

Evaluation von Adaptivität im E-Learning

Qualitätssicherung und Qualitätskontrolle einer multimedialen Lernumgebung in der Psychopathologieausbildung

Abhandlung
zur Erlangung der Doktorwürde
der Philosophischen Fakultät
der Universität Zürich

vorgelegt von
Matthias Häne
von Kirchberg / SG

Angenommen im Herbstsemester 2008 auf Antrag von
Prof. Dr. Damian Läge und
Prof. Dr. Wolfgang Marx

Zürich 2008

Inhaltsverzeichnis

| | | |
|-------|---|-----|
| I. | Einleitung | 1 |
| I.I | Einführung und Problemstellung | 2 |
| I.II | Konzept und Gesamtüberblick..... | 4 |
| I.III | Literaturverzeichnis | 9 |
| 1. | Die Lernumgebung | 10 |
| 1.1 | Inhaltsgebiet | 12 |
| 1.2 | Didaktische und methodische Ansätze | 14 |
| 1.3 | Lerninhalte | 22 |
| 1.4 | Lernerfreundliche Ausgestaltung..... | 26 |
| 1.5 | Der Lernprozess in PTO | 27 |
| 1.6 | Adaptives Lernen mit PTO | 28 |
| 1.7 | Literaturverzeichnis | 37 |
| 2. | Theoretische Grundlagen | 39 |
| 2.1 | Grundlagen der Evaluation | 41 |
| 2.2 | Evaluation im E-Learning | 61 |
| 2.3 | Literaturverzeichnis | 73 |
| 3. | Evaluationskonzept | 76 |
| 3.1 | Einleitung | 78 |
| 3.2 | Chancen der Adaptivität im E-Learning | 78 |
| 3.3 | Das Fallbeispiel: Die E-Learning-Umgebung PTO | 79 |
| 3.4 | Evaluationsplan..... | 82 |
| 3.5 | Schlussbemerkung | 85 |
| 3.6 | Literaturverzeichnis | 86 |
| 4. | Gesamtevaluation..... | 87 |
| 4.1 | Grundlagen der Evaluation von PTO..... | 89 |
| 4.2 | Evaluationsmethoden und Resultate | 94 |
| 4.3 | Diskussion..... | 114 |
| 4.4 | Literaturverzeichnis | 116 |

| | | |
|-----|--|-----|
| 5. | Adaptivitätsstudie | 117 |
| 5.1 | Einleitung..... | 119 |
| 5.2 | Methode | 121 |
| 5.3 | Fragestellung 1: Wissensmessung mit Kognitiven Karten | 126 |
| 5.4 | Fragestellung 2: Erfassung des Lernerfolgs..... | 129 |
| 5.5 | Fragestellung 3: Wirkung des Adaptiven Tutoriellen Systems | 134 |
| 5.6 | Lernprozess auf Individualebene | 142 |
| 5.7 | Diskussion..... | 144 |
| 5.8 | Literaturverzeichnis | 146 |
| 6. | Studie zur Lerneffizienz und Lerneffektivität..... | 147 |
| 6.1 | Theoretische Grundlagen..... | 149 |
| 6.2 | Experiment 1: Bild/Text-Kombination | 154 |
| 6.3 | Experiment 2: Pop-up-Glossar..... | 159 |
| 6.4 | Experiment 3: Speed-Reading | 162 |
| 6.5 | Diskussion..... | 167 |
| 6.6 | Literaturverzeichnis | 170 |

Zusammenfassung

Die Entwicklung der letzten Jahre im Bereich der Informations- und Kommunikationstechnologie hat zu neuen Lehr- und Lernformen geführt. Im Bildungssektor resultierte daraus eine intensive institutionelle Förderung und Verbreitung von E-Learning. Insbesondere auf Hochschulebene wurden die Vorteile computergestützten Lernens – individuelles, selbstorganisiertes, effizientes Lernen bei gleichzeitiger Entlastung der Hörsäle – früh erkannt. In der Folge wurde eine Vielzahl an Projekten mit dem Ziel multimedialer Wissensvermittlung initiiert. Die Erfahrung zeigte, dass der gewünschte Erfolg stark davon abhängig war, ob das entwickelte Produkt den Ansprüchen universitärer Bildung genügte. Die Frage nach der Qualität des Bildungsangebots rückte damit ins Zentrum des Interesses. Bei der Entwicklung der webbasierten Lernumgebung „Psychopathology Taught Online“ (PTO) stellte sich die Qualitätsfrage ebenfalls. Daher wurde im Sinne eines umfassenden Qualitätsverständnisses eine permanente Evaluationsstelle im Entwicklerteam geschaffen. Die systematische und konsequente Evaluation von PTO sollte die Sicherung und Kontrolle von Qualität für sämtliche relevanten Aspekte der Lernumgebung und bezogen auf den gesamten Entwicklungsprozess gewährleisten. Die Evaluation war umso bedeutender, da PTO neben den eigentlichen Lerninhalten über ein Adaptives Tutorielles System verfügt, welches die Lernenden bei der Gestaltung ihres Lernprozesses unterstützen soll. Dieses System basiert auf einer innovativen Verfahrenslage, welche sich unter Praxisbedingungen noch zu bewähren hatte. Der Komplexität des Gegenstands entsprechend, umfasste das Repertoire an Evaluationsmassnahmen verschiedenste methodische Ansätze: pragmatisch ausgerichtete Benutzertests, heuristische Evaluationen, Akzeptanzbefragungen der am Lernprozess beteiligten Personen, empirische Studien zur Überprüfung von Lerneffizienz und Lernerfolg einzelner in PTO implementierter didaktischer Elemente sowie ein Lernexperiment zur Untersuchung des Adaptivitätsmechanismus. Diese Erhebungen lieferten das empirische Datenmaterial, auf dessen Grundlage die Qualität von PTO adäquat eingeschätzt und bewertet sowie laufend verbessert und gemäss den definierten Zielstellungen und Ansprüchen optimiert werden konnte.

Abstract

In recent years, developments in the area of information and communication technology have lead to new forms of teaching and learning which have prompted an intense institutional furtherance and diffusion of E-Learning. Universities in particular soon recognized the advantages computer based learning could offer: individual, self-organised, efficient learning and less crowded lecture halls. Consequently, a vast array of projects examining different dimensions of multimedia based knowledge transfer emerged. Experience

showed that E-Learning success was highly dependent on the degree to which the developed product met the demands of college education. The focus of attention has therefore shifted to the quality of the training courses. In the development of the web-based learning environment “Psychopathology Taught Online” (PTO) the question of the quality of the training course was therefore of prime importance. In order to manage quality a constant process of evaluation was undertaken. The systematic and consistent evaluation of PTO was developed to control and ensure the quality of all relevant aspects of the learning environment and the process development as a whole. Evaluation was all the more important since PTO was endowed with an Adaptive Tutorial System that supported learners in the organisation of their learning processes. The system was based on an innovative approach that had yet to prove itself in the working environment. The complexity of the matter demanded that an array of methods be used: pragmatically orientated user-tests, heuristic evaluation, acceptance surveys for all who took part in the learning process, empirical studies to test the learning efficiency and learning success of several PTO implemented didactical elements, and a learning experiment to check the functioning of the adaptivity mechanism. These surveys provided the empirical data needed to adequately assess and evaluate the quality of PTO. That way the learning environment could continuously be improved and optimised in accordance with the defined goals and demands.

Einleitung

I.1 Einführung und Problemstellung

Informations- und Kommunikationstechnologien gewinnen zunehmend an Bedeutung und durchdringen immer mehr Bereiche gesellschaftlichen Lebens. In Wissenschaft, Wirtschaft, Politik und Unterhaltung ist der Einfluss der neuen Medien in den letzten Jahren stark gestiegen und schliesst auch die Bereiche Bildung und Weiterbildung mit ein. Die rasante Entwicklung sowie beschleunigte Innovationszyklen der Informations- und Kommunikationstechnologien führten zu neuen Lehr- und Lernformen. Der Wandel zur heutigen Wissensgesellschaft und mit ihr der Anspruch, das verfügbare Wissen über die Wissensmedien immer mehr Menschen immer schneller verfügbar zu machen, hat diesen Prozess zusätzlich gefördert (Meister, 2004). Der medienbasierte Wissensaustausch hat im Zusammenhang mit dieser Entwicklung zu einer neuen Lernkultur geführt mit neuen Anforderungen für die im Bildungsprozess involvierten Personen.

Im Bildungsbereich hat sich dabei der E-Learning-Begriff etabliert. E-Learning als Oberkategorie umfasst sämtliche Formen elektronischer Wissensvermittlung, bei denen Lerninhalte mit Hilfe von Computern multimedial dargeboten werden¹. Dazu gehören beispielsweise das Computer-Based-Training (CBT), bei dem die Lernprogramme lokal auf dem Computer des Benutzers verfügbar sind, oder das Web-Based-Training (WBT), welches das Lernen mittels auf Web-Technologie (Intranet und Internet) basierenden Anwendungen beinhaltet. Im Gegensatz zum traditionellen Lernen zeichnet sich E-Learning durch verschiedene Besonderheiten aus (Arnold, 2006; Baumgartner, 1997; Handke, 2003): Eine wesentliche Eigenschaft ist die Möglichkeit der Interaktion zwischen Benutzer und System. Dieser ist nicht bloss Rezipient, sondern greift aktiv in den Lernprozess ein. Weiter ist für E-Learning eine hohe Flexibilität charakteristisch, da es weder orts- noch zeitgebunden erfolgt. Neben selbstorganisiertem, individualisiertem Lernen fördert es verschiedene Formen von Kooperationen, was einen intensiven Wissensaustausch ermöglicht. Im Bildungsbetrieb findet zudem vermehrt ein Konzept Anwendung, welches in einer Form hybriden Lernens die Vorteile von E-Learning mit denen einer Präsenzvorlesung kombiniert. Dieser sogenannte Blended-Learning-Ansatz propagiert die pädagogisch sinnvolle Einbettung von E-Learning in Lernarrangements, wovon man sich einen grösseren Lernerfolg verspricht. E-Learning führt weiter zu einer Erweiterung des didaktischen und medialen Repertoires der Lehrenden und ermöglicht problemorientierte Lernformen (Schaper, 2005). Daraus resultiert ein neues Rollenverständnis sowohl für die Lehrpersonen als auch für die Lernenden. Erstere sehen sich je nach Lernszenario mit neuen Anforderungen konfrontiert, welche den Erwerb zusätzlicher Kompetenzen nötig machen. Letztere übernehmen mehr Eigenverantwortung im Lernprozess und gestalten

¹ Das „E“ in E-Learning steht für „electronic“ und bezieht sich auf die elektronische Informationsverarbeitung und die elektronisch gestützte Kommunikation beim Lernen.

diesen aktiver. Zusammenfassend lässt sich sagen, dass der Bildungsbereich insbesondere auf Hochschulebene (in diesen Zusammenhang gehören die Schlagworte Virtueller Campus oder Virtuelle Hochschule) durch das E-Learning starke Veränderungen erlebt. Auf der einen Seite eröffnen sich dadurch neue Möglichkeiten und Chancen. Auf der anderen Seite sind mit dieser Entwicklung jedoch auch Risiken verbunden. Auf Grund der jungen Entwicklung ist das Wissen über die Möglichkeiten und Grenzen von E-Learning noch fragmentarisch.

Gegen Ende der 1990er Jahre setzte eine Phase der intensiven Förderung von Projekten im Bereich von Informationstechnologie und E-Learning ein. In der Schweiz ist in diesem Zusammenhang vor allem der Swiss Virtual Campus (SVC) zu nennen, welcher webbasiertes Lernen auf Hochschulebene unterstützte. In der Folge wurden eine Vielzahl multimediasbasierter Lerninhalte, Kurse oder vollständige online-Studiengänge produziert. Die Förderprogramme sind zwischenzeitlich mehrheitlich ausgelaufen. Es lässt sich feststellen, dass es einem beträchtlichen Teil der Projekte nicht gelungen ist, sich längerfristig zu bewähren. Dies hängt unter anderem damit zusammen, dass die befristet angestellten Mitarbeiter nach Abschluss der Entwicklung die Projekte teilweise verliessen. Mit ihnen ging aber auch das projektspezifische Know-how verloren, was eine Fortführung erschwerte und mit Qualitätseinbussen einherging. Oftmals sind auch die mangelnde Qualität des didaktischen Konzepts oder der technischen Umsetzung sowie die mangelnde Ausrichtung am Benutzer für das Scheitern verantwortlich. Gefragt sind heutzutage daher langfristig orientierte Strategien und Lösungen, welche eine nachhaltige Projektentwicklung sicherstellen und welche die qualitativen Ansprüche der verschiedenen Anspruchsgruppen erfüllen.

Grundsätzlich lässt sich festhalten, dass der erfolgreiche Einsatz von E-Learning stark von der Qualität der Systeme abhängig ist. Was aber Qualität ist und was sie im Zusammenhang mit E-Learning bedeutet, ist eine Frage, welche sehr unterschiedlich beantwortet werden kann. Nach einer ersten Phase, in der hauptsächlich technische Eigenschaften im Vordergrund standen, besteht mittlerweile zumindest Konsens darüber, dass die Qualität medienbasierten Lernens ein vielschichtiges Konstrukt ist, dem eine eindimensionale Betrachtungsweise nicht gerecht wird. Der Frage, was Qualität überhaupt ist, folgt die Frage, wie diese gesichert werden kann (Ehlers, 2002). Qualitätssicherung im E-Learning kann generell auf dieselben Konzepte, Modelle und Methoden zurückgreifen, welche sich bereits in anderen Bereichen bewährt haben. In den wenigsten Fällen können diese jedoch einfach übernommen werden, sondern müssen den Anforderungen des Kontexts gemäss angepasst und modifiziert werden. Die Disziplin schliesslich, welche sich mit der Untersuchung von Qualität beschäftigt und die die Grundlage zur Sicherung und Kontrolle von Qualität liefert, ist die Evaluation. Diese dient der Erhebung von Daten, welche eine adäquate Einschätzung von Bildungsangeboten unter dem jeweils interessierenden Beurteilungsaspekt ermöglichen (Tergan, 2000).

Im Rahmen der Entwicklung der webbasierten Lernumgebung Psychopathology Taught Online (PTO) stellten sich diese Fragen nach Qualität und deren Gewährleistung ebenfalls. PTO bietet eine strukturierte Einführung in die Phänomenologie psychopathologischer Störungsbilder sowie einen Einblick in ausgewählte Themen der Psychopathologie und soll ergänzend zum universitären Lehrangebot als digitales Curriculum eingesetzt werden. Daneben verfügt es über ein Adaptives Tutorielles System (ATS). Den Lernenden werden konkrete, an ihrem momentanen Wissensstand angepasste Lernempfehlungen gegeben, was den Lernprozess individualisierter und effizienter gestaltet.

Von Beginn an wurde Qualität und deren Sicherung als wesentlicher Faktor für das Gelingen des Projekts erkannt. Aus diesem Grund wurde innerhalb des Projektteams eine permanente Evaluationsstelle geschaffen, welche sich spezifisch mit Fragen nach Qualität und Optimierung auseinander setzte. Die Qualität von PTO sollte während der gesamten Entwicklung sichergestellt werden, woraus sich der Anspruch nach einem umfassenden Qualitätsmanagement ergab. Im Sinne dieses Qualitätsverständnisses wurde es als wesentlich angesehen, dass die Lernumgebung in sämtlichen Entwicklungsphasen systematisch und unter Berücksichtigung der relevanten Aspekte getestet und evaluiert wurde. Diese Analysen dienten verschiedenen Zwecken: Als Kontroll- und Entscheidungsfunktion sollten die Evaluationsmassnahmen die Datengrundlage generieren, anhand derer der Entwicklungsprozess überwacht und bei Bedarf optimiert werden konnte. Da PTO vom SVC finanziert wurde, bestand dieser Institution gegenüber eine Rechenschaftspflicht. Als strategisch-politische Funktion sollten die Evaluationsdaten daher der Überwachung der Projektperformance dienen. Zudem sollten sie den potenziellen Anwendern den Sinn und Nutzen von PTO aufzeigen und hatten damit Anreizfunktion. Evaluation als Erkenntnisfunktion schliesslich sollte die Effekte aufdecken, welche mit den in PTO verwendeten mediendidaktischen Mitteln erzielt werden können.

Den Rahmen der vorliegenden Arbeit bildet die Frage nach der Qualität im Kontext von PTO. Somit werden sämtliche Aktivitäten und Massnahmen beschrieben, welche im Zusammenhang mit der Evaluation von PTO anfielen. Die theoretischen Grundlagen der Evaluation werden dargelegt und die daraus abgeleiteten Evaluationskonzepte vorgestellt. Weiter werden Idee, Aufbau und didaktisches Konzept der Lernumgebung präsentiert, auf welche sich die Evaluationsbestrebungen beziehen. Zudem wird anhand von empirischen Studien der Lernprozess sowie die Wirkung verschiedener im Kontext von PTO entwickelter didaktischer Elemente evaluiert.

I.II Konzept und Gesamtüberblick

Die folgenden Ausführungen geben eine summarische Übersicht über den Prozess der Evaluation von PTO sowie über die Ergebnisse, welche dabei erzielt wurden. Für eine

ausführliche Darstellung und Würdigung der Resultate sei auf die einzelnen Kapitel verwiesen.

Mit den theoretischen Ausführungen zur Evaluation wurde das Fundament gelegt, auf das sich die Aktivitäten und Massnahmen im Rahmen der Arbeit stützen und welche die systematische und theoriegeleitete Evaluation von PTO erst ermöglichten. Dabei fanden der Thematik dieser Arbeit entsprechend Aspekte der Evaluation im Bereich E-Learning besondere Berücksichtigung. Anschliessend wurde ein erstes Evaluationskonzept formuliert. Der Konzeptentwurf stand dabei ganz zu Beginn des Evaluationsprozesses von PTO. In der Folge wurde dieser noch weiter differenziert und elaboriert. Das finale Konzept wurde aus den im Theorieteil formulierten Grundlagen hergeleitet. Kern des Konzepts bildet eine Evaluationsmatrix, bestehend aus den Achsen Evaluationsart (formative und summative Evaluation) und Evaluationsebene (Produkt-, Reaktions- und Lernebene). Damit war sichergestellt, dass die Evaluation sowohl Aspekte der Qualitätssicherung (formative Evaluation) als auch solche der Qualitätskontrolle (summative Evaluation) beinhaltet und sich auf den gesamten Entwicklungsprozess bezieht. Im Sinne eines multipolaren Qualitätsansatzes wurden sämtliche Ebenen untersucht, welche für den erfolgreichen Einsatz von PTO konstituierend sind: zunächst die eigentliche Lernumgebung als System (Produktebene), dann die Akzeptanz auf Seiten der Nutzer (Reaktionsebene) und schliesslich die Wirkung in Form des Lernerfolgs, welcher sich bei der Bearbeitung von PTO einstellt (Lernebene).

Auf Produktebene bezog sich die Evaluation auf einzelne Komponenten des Systems, welche als besonders relevant taxiert wurden. In einem ersten Schritt wurden in diesem Zusammenhang Aspekte der Benutzungstauglichkeit überprüft. Das Ziel bestand in einer reibungslosen und barrierefreien Interaktion zwischen Lernenden und System, welche eine notwendige Voraussetzung für effizientes Lernen darstellt. Die Evaluation lieferte die Datenbasis zur entsprechenden Anpassung des Systems. Zudem wurde in einem Lernexperiment der Adaptivitätsmechanismus getestet, welcher in PTO erstmalig zum Einsatz kommt und welcher der Realisierung eines zielgerichteten und effektiven Lernprozesses dient. Das Adaptive Tutorielle System basiert auf einer innovativen Verfahrenslogik, welche sich in der Praxis noch bewähren musste. Die technischen Möglichkeiten einer multimedialen Lernumgebung nutzend, wurden im Kontext von PTO verschiedene didaktische Elemente entwickelt. Diese sollten als Lernhilfen unterstützend wirken und den Lernprozess insgesamt einfacher und effizienter gestalten. Der Nutzen dieser didaktischen Elemente wurde in Bezug auf Lerneffizienz und Lernerfolg im Rahmen einer empirischen Studie getestet.

Auf der Reaktionsebene stand die Akzeptanz der Nutzer von PTO im Fokus. Diese stellt eines der zentralen Elemente einer erfolgreichen Einführung eines E-Learningangebots dar. Entsprechend wurden zu deren Erhebung an verschiedenen Zeitpunkten im Entwick-

lungsverlauf Akzeptanzbefragungen durchgeführt. Auf der Lernebene interessierte besonders der Lernerfolg, welcher sich bei der Arbeit mit PTO erzielen lässt. Um diesen zu messen, wurde ein Lernexperiment durchgeführt. Zu unterschiedlichen Zeitpunkten im Lernprozess wurde dazu das Wissen zu den in PTO vermittelten Störungsbildern auf Ebene des strukturellen wie auch des deklarativen Wissens erfasst.

Anhand der Evaluationsmatrix liess sich PTO systematisch testen und evaluieren. Dies lieferte das Datenmaterial, auf dessen Grundlage Aussagen bezüglich der Qualität der Lernumgebung gemacht werden konnten. Die zur Evaluation eingesetzten Methoden umfassten ein breites Spektrum empirischer Forschungstätigkeit (heuristische Evaluation, Usability-Testing, Befragungen, empirische Studien, Lernexperiment). Wo sich Diskrepanzen zwischen Anspruchsniveau und tatsächlichem Zustand zeigten, konnten entsprechend Massnahmen zur Verbesserung und Optimierung eingeleitet werden.

Die vorliegende Arbeit setzt sich aus verschiedenen Einzelbeiträgen zusammen, wobei jeder dieser Beiträge einen bestimmten Aspekt im Evaluationsprozess von PTO beschreibt. Die Beiträge sind, um diesem Prozess Rechnung zu tragen und um das zielgerichtete Zusammenspiel der verschiedenen Aktivitäten zu verdeutlichen, gemäss ihrer Stellung im chronologischen Ablauf des Evaluationsprozesses angeordnet. Die einzelnen Beiträge wurden so verfasst, dass sie auch für sich gelesen werden können. Entsprechend liess sich ein gewisses Mass an Redundanz nicht vermeiden.

Die Lernumgebung PTO

Kapitel 1 vermittelt in knapper Form einen Überblick über sämtliche relevanten Aspekte des Projekts PTO. So wird der Inhaltsbereich von PTO sowie der methodisch-didaktische Ansatz vorgestellt, welcher im Hinblick auf die zu erfüllenden Anforderungen gewählt wurde. Weiter werden die Struktur und der Aufbau von PTO erklärt, die Lerninhalte dargelegt sowie der idealtypische Lernprozess beschrieben wie er für PTO konzipiert wurde. Schliesslich wird das in PTO implementierte Adaptive Tutorielle System vorgestellt.

Theoretische Grundlagen

Kapitel 2 befasst sich mit den theoretischen Grundlagen der Evaluation und der Evaluationsforschung. Es wird eine Annäherung an den Evaluationsbegriff unternommen sowie auf wichtige Aspekte der Evaluation im historischen Kontext eingegangen. Weiter werden die verschiedenen Arten und Typen von Evaluation dargelegt, deren Funktion erläutert und die Anwendung der Methoden im Entwicklungsprozess diskutiert. Schliesslich wird spezifisch auf Evaluation im Kontext von E-Learning eingegangen. Diese ist im Zusammenhang mit der Evaluation von PTO von zentraler Bedeutung.

Evaluationskonzept

Kapitel 3 beinhaltet einen Artikel, der als Proceeding-Beitrag im Rahmen der Jahrestagung der Gesellschaft für Medien in der Wissenschaft (GMW) publiziert wurde². Ausgangslage bildete die Finanzierung des E-Learning-Projekts PTO durch den Swiss Virtual Campus (SVC). Für dieses Projekt galt es, im Sinne eines umfassenden Qualitätsmanagements die Sicherung und Kontrolle der Qualität zu gewährleisten. Der Beitrag legt daher ein erstes Konzept für die Evaluation von PTO fest.

Gesamtevaluation

In Kapitel 4 wird das definitive Konzept zur Evaluation auf Basis der in Kapitel 2 eingeführten theoretischen Grundlagen hergeleitet. Als Ergebnis wird ein Handlungsrahmen zur zielgerichteten, systematischen Evaluation von PTO eingeführt. Dieser gewährleistete die formative (qualitätssichernde) sowie die summative (qualitätskontrollierende) Evaluation der für das Projekt relevanten Aspekte. Die eingesetzten Evaluationsmassnahmen bezogen sich dabei auf die Ebenen Aufbau und Struktur der Lernumgebung (Produktebene), die Akzeptanz auf Seiten der Nutzer (Reaktionsebene) sowie die Effekte, welche sich mit der Bearbeitung von PTO erzielen lassen (Lernebene). Dazu werden die verwendeten Evaluationsmethoden beschrieben und ausgewählte Resultate präsentiert.

Adaptivitätsstudie

Kapitel 5 beschreibt eine empirische Studie in Form eines Lernexperiments, welches der Untersuchung des in PTO implementierten Adaptiven Tutoriellen Systems diene und insgesamt drei Ziele verfolgte: Erstens sollte untersucht werden, ob sich mittels Kognitiver Karten Wissen messen lässt³. Diese Untersuchung war von Bedeutung, da Kognitive Karten in PTO zur Wissensdiagnose herangezogen werden. Als externes Kriterium zur Überprüfung der Validität dieser Karten wurden Multiple-Choice-Tests durchgeführt. Es zeigte sich ein signifikanter Zusammenhang zwischen der Güte des strukturellen Wissens und dem Abschneiden im Multiple-Choice-Test. Damit liess sich bestätigen, dass die relationale Wissensdiagnose und die darauf basierenden Kognitiven Karten ein valides Instrument zur Wissensmessung darstellen.

Zweitens sollte der Lernerfolg erfasst werden, welcher bei der Arbeit mit PTO erzielt werden kann. Das Wissen über die Störungsbilder wurde zum einen mittels Multiple-Choice-Test auf der Ebene des deklarativen Wissens erhoben. Zum anderen wurde dieses Wissen auch mittels Kognitiver Karten auf struktureller Ebene erfasst. Dies geschah zu

² „Adaptivität und deren Evaluation im E-Learning“, Häne, Streule, Egli, Oberholzer und Läge, 2006.

³ Kognitive Karten basieren auf relationalen Daten und repräsentieren das strukturelle Wissen zu einem bestimmten Inhaltsgebiet.

verschiedenen Zeitpunkten im Lernprozess. Durch den Vergleich der Ergebnisse zu den unterschiedlichen Erhebungszeitpunkten liess sich auf den Lernerfolg schliessen. Hier zeigte sich im Lernprozess eine signifikante Verbesserung sowohl auf der Ebene des deklarativen Wissens (Resultate im Multiple-Choice-Test) als auch auf Ebene des strukturellen Wissens (Güte der Kognitiven Karten).

Drittens sollte das Funktionieren des Adaptivitätsmechanismus überprüft werden, welcher in PTO zur Anwendung kommt. Durch Vergleich der mit der Wissensdiagnose erhobenen Lernerstruktur mit einer Expertenstruktur sind Wissensmängel detektierbar. In der Folge werden spezifische, individuell ausgerichtete Lernempfehlungen abgegeben. Es wurde überprüft, wie sich die Wissensstrukturen der Lerner durch die Bearbeitung der Lernempfehlungen änderten. Die repetierten Störungsbilder verbesserten sich bezüglich ihrer Position in der Wissensstruktur deutlich. Auf individueller Ebene konnte damit das Wirken des Adaptiven Tutoriellen Systems aufgezeigt werden. Auf Stichprobenebene unterstützen die quantitativen Daten diesen Befund.

Studie zur Lerneffizienz und Lerneffektivität

Das Kapitel 6 beinhaltet eine Studie zur Wirkung verschiedener didaktischer Elemente, welche in PTO zum Einsatz kommen. Die technischen Möglichkeiten des E-Learnings führen zu einer Erweiterung des Repertoires an didaktischen Mitteln. Diese sollen helfen, die Lerninhalte so aufzubereiten, dass sie für die Lerner mit weniger Aufwand gelernt und besser memoriert werden können. Drei im Kontext von PTO entwickelte didaktische Elemente wurden bezüglich Lerneffektivität (Lernerfolg) und Lerneffizienz untersucht. Zum einen handelt es sich dabei um eine interaktive Grafik zur Angstsymptomatik, bei welcher der Lerner als Lernhilfe einzelne Elemente ein- und wieder ausblenden kann. Ein weiteres didaktisches Element ist der Pop-up-Glossar. In PTO sind Fremd- und Fachwörter auf einfache Art und Weise abrufbar. Wird der Mauszeiger über einen solchen Begriff platziert, wird die Definition des Begriffs im Mouse-over angezeigt. Dadurch lässt sich bei Bedarf kontextspezifisch auf Informationen zurückgreifen, wodurch Unterbrechungen und Störungen im Lernprozess vermieden werden. Als drittes Element wurde das Speed-Reading evaluiert. Diese Funktion ermöglicht es dem Lerner, mittels Knopfdruck die relevantesten Informationen markiert angezeigt zu erhalten. Dies soll helfen, den Lernprozess effizienter und effektiver zu gestalten. Insbesondere auf Ebene der Lerneffizienz zeigt sich deutlich, dass sich mit den Mitteln des multimedial basierten Wissenstransfers Lerninhalte ohne qualitative Einbussen innert kürzerer Frist vermitteln lassen.

I.III Literaturverzeichnis

- Arnold, P. (2006). Qualitätsentwicklung im E-Learning – Ansätze, Herausforderungen und Perspektiven. In B. Schwarz & D. Behrmann (Hrsg.), *Integratives Qualitätsmanagement – Perspektiven und Praxis der Organisations- und Qualitätsentwicklung in der Weiterbildung* (S. 79 – 114). Bielefeld: Bertelsmann.
- Baumgartner, P. (1997). Evaluation vernetzten Lernens: 4 Thesen. In H. Simon (Hrsg.), *Virtueller Campus. Forschung und Entwicklung für neues Lehren und Lernen* (S. 131-146). Münster: Waxmann.
- Ehlers, U. (2002). *Qualität im E-Learning: Der Lernende als Grundkategorie bei der Qualitätssicherung* [On-line]. Available: <http://www.medienpaed.com/02-1/ehlers1.pdf>
- Handke, J. (2003). *Multimedia im Internet*. München: Oldenbourg.
- Häne, M., Streule, R., Egli, S., Oberholzer, R. & Läge, D. (2006). Adaptivität und deren Evaluation im E-Learning - Das Fallbeispiel "Psychopathology Taught Online" (PTO). In E. Seiler Schiedt, S. Kälin & C. Sengstad (Hrsg.), *E-Learning - alltagstaugliche Innovation? Medien in der Wissenschaft, Band 38* (S. 296-305). Münster: Waxmann.
- Meister, D. (2004). *Online-Lernen und Weiterbildung*. Wiesbaden: Verlag für Sozialwissenschaften.
- Schaper, N. (2005) *Evaluation: Wissen über die Wirksamkeit von eLearning*. [On-line]. Available: <http://miami.uni-muenster.de/servlets/DerivateServlet/Derivate-2496/VortragSchaper.pdf>
- Tergan, S.O. (2000). Grundlagen der Evaluation. In P. Schenkel, S.O. Tergan & A. Lottmann (Hrsg.), *Qualitätsbeurteilung multimedialer Lern- und Informationssysteme* (S. 22–51). Nürnberg: BW Bildung und Wissen.

1.

Die Lernumgebung

PTO

Abstract

Dieses Kapitel vermittelt in knapper Form einen Überblick über sämtliche relevanten Aspekte des Projekts PTO. So wird der Inhaltsbereich von PTO sowie der methodisch-didaktische Ansatz vorgestellt, welcher im Hinblick auf die zu erfüllenden Anforderungen gewählt wurde. Weiter werden die Struktur und der Aufbau von PTO erklärt, die Lerninhalte dargelegt sowie der idealtypische Lernprozess beschrieben wie er für PTO konzipiert wurde. Schliesslich wird das in PTO implementierte Adaptive Tutorielle System vorgestellt, welches den Lernern zu einem effizienten und auf die eigenen Bedürfnisse angepassten Lernprozess verhelfen soll.

1.1 Inhaltsgebiet

Mit dem Titel "Psychopathology Taught Online" (kurz: PTO) wird klar auf einen bestimmten Inhaltsbereich verwiesen. Im Folgenden wird dieser Inhaltsbereich definiert und seine Auswahl als Thema für ein virtuelles Lernangebot begründet.

1.1.1 Psychopathologie

PTO vermittelt Basiskenntnisse im Fach Psychopathologie. Für das Verständnis der curricularen Einbettung einerseits und die Definition der Zielgruppe andererseits ist es hilfreich, diesen Fachbereich und seine Schnittstellen zu definieren.

Psychopathologie ist ein Teilbereich der Psychiatrie, das heisst einer medizinischen Fachrichtung. Wörtlich übersetzt bedeutet der Begriff "Lehre von den Leiden der Seele" oder "Lehre vom seelisch Abnormen", was an sich grossen Interpretationsspielraum zulässt. Im engeren Sinn beschäftigt sich das Fach jedoch mit der *Beschreibung* abnormen Erlebens, Befindens und Verhaltens (Möller, Laux & Deister, 2005). Psychopathologie konzentriert sich also auf deskriptive Aspekte psychischer Störungen, obwohl die oben genannten Autoren auch die verstehende Psychopathologie erwähnen, die nach "den inneren erlebens- und biographiebezogenen Zusammenhängen der psychischen Störungen" fragt (S. 1). Der Schwerpunkt von PTO liegt dagegen, in der Nomenklatur derselben Autoren, auf der klassifikatorischen Psychopathologie, also auf der Beschreibung, Benennung und Ordnung im Hinblick auf die Klassifikation psychischer Störungen.

Psychiatrie und mit ihr die Psychopathologie haben einen Überschneidungsbereich mit der Psychologie als Geisteswissenschaft. Zwar beschäftigt sich die Psychologie grösstenteils mit "normalem" Erleben und Verhalten, sie beinhaltet aber an praktisch jeder Institution, die eine umfassende Ausbildung in Psychologie anbietet, auch den Fachbereich Klinische Psychologie. Dieser befasst sich "mit psychischen Störungen und den psychischen Aspekten somatischer Krankheiten respektive Störungen. Dazu gehören unter anderem die Themen Ätiologie/Bedingungsanalyse, Klassifikation, Diagnostik, Epidemiologie, Intervention" (Baumann & Perrez, 2005, S. 32). Die Psychopathologie dient in diesem Zusammenhang der Beschreibung psychischer Störungen.

In diesem Sinn ist Psychopathologie für Psychiatrie und Psychologie einerseits eine humanwissenschaftliche Methodenlehre, andererseits umfasst sie auch eine definierte Expertensprache zur Erfassung von abnormen seelischen Zuständen und psychischen Störungen/Krankheiten.

PTO vermittelt neben ausgewählten, spezifischen Themen im Bereich Psychopathologie vor allem Wissen über die Erscheinungsbilder einer grösseren Anzahl von psychischen Störungen des Erwachsenenalters. Diesen Störungen begegnen Praktikerinnen und Praktiker in medizinischen, psychologischen, heilenden oder sozialen Berufen mit grosser

Wahrscheinlichkeit. Unter Erscheinungsbild wird im Folgenden das verstanden, was an einem Patienten für Aussenstehende beobachtbar und für den Patienten selbst erlebbar ist. Mit anderen Worten: Ein grosser Teil von PTO bringt den Lernenden die Symptomatik psychischer Störungen nahe. Benutzer und Benutzerinnen lernen, welche Symptome und Syndrome bekannte Störungen charakterisieren und diese Symptomatik in konkreten Fallbeispielen zu erkennen. Andere Aspekte der Psychiatrie und klinischen Psychologie wie Ätiologie, Epidemiologie oder Therapie werden grösstenteils ausgeklammert.

1.1.2 Gründe für die inhaltliche Fokussierung

Der enge inhaltliche Fokus von PTO lässt sich mehrfach begründen: Zum einen wurde Psychopathologie weiter oben als beschreibende Wissenschaft definiert. Die erwähnten weiteren Bereiche von Psychiatrie und Klinischer Psychologie gehören im engeren Sinn nicht in dieses Gebiet. Zudem beruhen gerade in der Klinischen Psychologie die Erklärungen für psychische Störungen und konsequenterweise auch deren Behandlungsformen auf unterschiedlichen theoretischen Sichtweisen (vor allem psychodynamisch, lerntheoretisch und kognitionspsychologisch). Diese abzudecken, hätte den Rahmen der Lernumgebung bei Weitem gesprengt.

Zweitens versteht sich PTO in erster Linie als ergänzendes Lernangebot, das als modularer Bestandteil in einen grösseren Ausbildungskontext eingebettet ist. Wie andere akademische Fachgebiete verwendet auch die Lehre in Fächern, bei denen Psychopathologie eine Teildisziplin darstellt, nach wie vor und zu einem grossen Teil konventionelle und bewährte Lernformen wie Präsenzveranstaltungen und das Studium traditioneller Lehrbücher. In solchen Gefässen findet etwa die Diskussion der oben erwähnten theoretischen Perspektiven ihren Platz.

Schliesslich sind die eher deklarativen Inhalte, zu denen die fallunabhängige Symptom- und Störungskunde gehört, aufgrund ihres repetitiven Charakters für Dozierende verständlicherweise weniger attraktiv als etwa individuelle Fallbesprechungen. Die Menge an Symptomen und Störungen muss ausserdem in Präsenzveranstaltungen recht zügig behandelt werden. So ist die Aneignung von psychopathologischem Basiswissen meist mit einem beträchtlichen Anteil an Selbststudium und Auswendiglernen verbunden. Da die Lernenden bei diesem Selbststudium weitgehend auf sich allein gestellt sind, macht es für Dozierende Sinn, diese grundlegenden Inhalte in ein strukturiertes E-Learning-Angebot auszulagern, um die Qualität dieses Teils des Lernprozesses sicherzustellen.

1.1.3 Herausforderungen

Diesen Argumenten zu Gunsten eines E-Learning-Angebots im Fach Psychopathologie stehen einige Herausforderungen gegenüber: Wie oben erwähnt, ist das Basiswissen über

die Erscheinungsbilder psychischer Störungen eher trockener Natur: Auch wenn die konkreten Erscheinungsformen in der klinischen Praxis äusserst spannend sein können, beschränkt sich ihre Beschreibung in Grundlagenbüchern meist auf Kriterienlisten (Kriterien sind Symptome sowie Angaben zur für eine Diagnose minimal erforderlichen Anzahl, Dauer und Schwere dieser Symptome).

E-Learning hat zudem im Fächerbereich rund um die Psychopathologie aus gutem Grund nur zu einem beschränkten Teil Einzug gehalten (meist in Form rudimentärer Angebote wie herunterladbare Skripts und Vorlesungsfolien): Es sind ausgesprochen textlastige Disziplinen. Viele Inhalte sind schwierig zu visualisieren, und die prozessualen Aspekte jener Wissenschaften, die sich der psychopathologischen Begrifflichkeit bedienen, lassen sich besser in Präsenzveranstaltungen vermitteln. Solche Prozesse bedürfen der unmittelbaren gedanklichen Modellierung durch die vermittelnde Fachperson und eines direkten individuellen Feedbacks, wenn die Lernenden diese nachvollziehen.

So steht PTO vor der Aufgabe, einen Wissensbereich zu vermitteln, der einerseits recht trocken wirkt und andererseits just solche Inhalte umfasst, die sich aufgrund ihrer eher verbalen Natur weniger für eine Umsetzung mittels E-Learning anbieten. Die folgenden Ausführungen schildern, wie PTO diesen Problemen begegnet.

1.2 Didaktische und methodische Ansätze

Im Folgenden werden erst die Begriffe „Lernen“ und „Didaktik“ definiert, bevor die didaktische Grundposition von PTO erläutert und die methodischen Konsequenzen daraus geschildert werden.

1.2.1 Lernen

Die folgenden Ausführungen weisen auf ein definitorisches Problem hin. Was man unter "Lernen" oder "gelernt haben" versteht, ist abhängig von der gewählten theoretischen Sichtweise. Deshalb soll an dieser Stelle festgehalten werden, welche Position hier vertreten wird.

Gemeinsam ist allen Definitionen, dass sie auf mittelfristig stabile Veränderungen abzielen und dass diese auf Erfahrung oder Übung des Individuums beruhen müssen und nicht etwa auf Reifung, Ermüdung oder die Wirkung chemischer Substanzen (Schermer, 1998). Die Veränderungen können äusserlich beobachtbar, also im weitesten Sinne sichtbare Verhaltensänderungen sein, wie dies beim klassischen und instrumentellen Konditionieren der Fall ist. Andere Sichtweisen orientieren sich hingegen an internen (mental)en Repräsentationen – sogenannten kognitiven Strukturen – und deren Aufbau oder Veränderungen. Der Ausdruck kognitiv ist vielfältig und umfasst nahezu alle intellektuellen Leistungen, von der Wahrnehmung über Denken und Problemlösen bis hin zur Bewusst-

heit. Er bezeichnet Vorgänge, "durch die ein Organismus Kenntnis von seiner Umwelt erlangt" (Edelmann, 2000, S. 114), charakterisiert aber auch interne Prozesse wie Urteilen, Schlussfolgern, Vorstellen usw. In Verbindung mit dem Begriff Strukturen – damit sind Elemente und die zwischen ihnen bestehenden Verbindungen gemeint – sind aber vor allem Wissensbestände und die hypothetische Form ihrer Speicherung angesprochen. Genauere Einsichten in die intern ablaufenden Prozesse und die Eigenschaften interner Repräsentationen gewann man vor allem mit den Ansätzen der Kognitionspsychologie ab der zweiten Hälfte des 20. Jahrhunderts. Anders als beim Behaviorismus müssen sich unter dieser Perspektive die Veränderungen nicht zwingend in äusserem Verhalten manifestieren. Im kognitionspsychologischen Paradigma wird Lernen mittlerweile interpretiert als "Gesamtheit der Prozesse, die mit der Aufnahme von Informationen, ihrer weiteren Verarbeitung und Speicherung im Gedächtnis sowie ihrer Anwendung in spezifischen Situationen mit Aufgabencharakter verbunden sind" (Seel, 2000, S. 18).

Die Informationsverarbeitungs-Perspektive wurde ausserdem zunehmend beeinflusst von der Vorstellung von Lernen als individuelle Konstruktion von Wissen. Obschon einige Jahrzehnte früher begründet, gewann Piagets (1970) genetische Erkenntnistheorie diesbezüglich deutlichen Einfluss. Dieser Theorie gemäss ist, was wir Wirklichkeit nennen, nie absolut, sondern nur mit Hilfe unserer "Erkenntnisinstrumente" erfahrbare, das heisst mit unseren Sinnen, der Motorik und den mentalen Prozessen und Strukturen. Jegliche Erfahrung der äusseren Welt ist keine veridikale Abbildung, keine virtuelle Kopie, sondern das Resultat eines Rekonstruktionsprozesses mit Hilfe dieser Instrumente (deshalb der Ausdruck Konstruktivismus für diese Auffassung der Erfahrung der Wirklichkeit). "Erkenntnisinstrumente" können dabei Bewegungen oder Handlungen, aber auch Denkprozesse oder begriffliche Wissensbestände sein. Piaget spricht von Schemata, bei Denkprozessen auch von Operationen.

Neue Gegenstände oder Konzepte erfassen und begreifen wir, indem wir die uns verfügbaren Instrumente (Schemata) darauf anwenden. Lassen sich die verfügbaren Schemata problemlos und sinnvoll auf den neuen Sachverhalt anwenden, spricht Piaget von Assimilation. So assimiliert ein Kleinkind einen Wollknäuel an sein einhändiges Greifschema, wenn es ihm gelingt, den Knäuel mit der Greifbewegung seiner Hand zu erfassen. Ebenso wird – auf begrifflicher Ebene – ein Schulkind den Ausdruck Konto problemlos an sein Konzept von "Sparen" assimilieren können. Ist der Knäuel für das Kleinkind aber zu gross, wird er sich dem Zugriff mit nur einer Hand widersetzen. Das Kind ist gezwungen, beide Hände einzusetzen, um den Knäuel zu halten. Es modifiziert sein ursprüngliches Bewegungsmuster und passt sein Verhalten dem neuen Gegenstand an; Piaget spricht von Akkomodation, denn nun passen sich der Organismus und seine Strukturen der äusseren Gegebenheit an. In ähnlicher Weise – aber bezogen auf die kognitive Ebene – wird ein Schüler sein Schema von Säugetieren korrigieren respektive neu definieren müssen, wenn er erfährt, dass Wale zu den Säugetieren gehören.

Ein ausgewogenes Verhältnis dieser beiden grundlegenden Prozesse – Assimilation und Akkomodation – erlauben es dem Individuum, erfolgreich mit der Umwelt zu interagieren und sich veränderten oder neuen Bedingungen anzupassen. Piaget spricht von Äquilibration. Seine Erkenntnistheorie liefert ein generelles Konzept des Begreifens der Wirklichkeit, das Piaget neben Kants Erkenntnislehre auch aus der intensiven Beobachtung von Kindern gewonnen hat. Dementsprechend ist die Theorie auch stark handlungsbezogen: Jegliches Denken entwickelt sich für Piaget aus der handelnden Interaktion mit der Umwelt. Man könnte seine Theorie in dem Sinne als "dynamischen Kantianismus" bezeichnen, als er Erkenntnisfähigkeiten, die Kant (Kritik der reinen Vernunft) noch als im Menschen "a priori", also ohne Erfahrung vorhanden voraussetzte, ebenfalls als Produkte der handelnden Auseinandersetzung mit der Umwelt versteht.

Assimilation und Akkomodation manifestieren sich jedoch auch auf kognitiver Ebene. Stossen wir auf neue gedankliche Konzepte, so versuchen wir zuerst, sie mit Hilfe der bereits vorhandenen Wissensgefüge zu verstehen – auch wenn das bereits Bestehende vielleicht unvollständig oder fehlerhaft ist. Solange sich daraus keine Schwierigkeiten ergeben, scheint der Assimilationsprozess geglückt. Stossen wir aber auf Widersprüche beim Umgang mit dem neuen Konzept (etwa in der Kommunikation darüber mit anderen Personen), so kommen wir nicht umhin, in Frage zu stellen, was uns bisher als richtig erschien. Wir sehen uns gezwungen, unsere bisherigen kognitiven Strukturen zu korrigieren, um das neue Konzept widerspruchsfrei in unseren Wissensbestand integrieren zu können.

Ob durch einfache Assimilation oder eher aufwändige Akkomodation – es ist meist dieses "Zur-Passung-Kommen", von dem die Rede ist, wenn jemand vom "Verstehen" eines Sachverhalts spricht. Diese Auffassung davon, was sich auf kognitiver Ebene beim Lernen und Verstehen in den Köpfen der Lernenden tut, kann nun auf zwei durchaus unterschiedliche Arten wegweisend für die Konzeption eines Lernangebots sein. Da sich daraus in der Konsequenz auch unterschiedliche Formen von Lernumgebungen ergeben, wird es darum gehen, sich entweder für eine Richtung zu entscheiden, oder aber einzelne Aspekte beider Richtungen gewinnbringend zu vereinen, während andere nicht in aller Konsequenz realisiert werden. PTO verfolgt den zweiten Ansatz.

1.2.2 Didaktik

Die folgenden Ausführungen beschäftigen sich mit Überlegungen zum Einsatz von PTO innerhalb eines fachspezifischen Curriculums. Dazu werden das Verständnis des Begriffs Didaktik innerhalb dieser Arbeit geklärt und Antworten auf didaktische Fragestellungen gegeben.

Didaktik. Wie Unterricht im Allgemeinen und eine konkrete Lernumgebung im Speziellen gestaltet werden sollen, ist abhängig von den Antworten auf verschiedene Fragen, mit denen sich die Didaktik beschäftigt. Auch wenn der Begriff oft verschieden ausgelegt wird, sind sich die meisten Autoren und Autorinnen darin einig, dass sich Didaktik mit der Theorie des Lehrens und Lernens (prioritär aber mit Ersterem) beschäftigt. Sofern man den Begriff, wie es im Folgenden der Fall sein wird, recht breit definiert, könnte man Didaktik auch als Theorie des Unterrichts (nicht nur des Unterrichtens) auffassen. Das kann eine generelle bildungstheoretische Ebene betreffen, z. B. im Sinne einer durch die gesellschaftlichen und politischen Verhältnisse bestimmten Auswahl von Inhalten, die durch die Institution Schule vermittelt werden sollen. Manchmal bezieht sich der Begriff aber auch viel konkreter auf Fragen nach dem Wie des Lehrens oder nach guten Voraussetzungen, die das Lernen unterstützen. Mit Ballstaedt (1997) könnte man Didaktik in diesem Sinn als "angewandte Disziplin zur Beeinflussung von Lernprozessen" bezeichnen (S. 12).

Diese Arbeit orientiert sich an der Vorstellung von Jank und Meyer (2002). Diesen Autoren gemäss befasst sich Didaktik ganz allgemein damit, "wer was von wem wann mit wem wo wie womit und wozu lernen soll" (S. 16). Didaktik wird dabei sehr weit definiert als "Theorie und Praxis des Lehrens und Lernens". Sie wird also keineswegs als rein theoretische Wissenschaft verstanden, sondern schliesst praktische Gesichtspunkte der Umsetzung im konkreten Unterricht mit ein. Die frühere Auffassung, derzufolge sich Didaktik primär um das Was kümmere, während die Methodik sich mit dem Wie des Unterrichtens befasse, wird somit abgelöst durch eine hierarchische Beziehung der beiden Gebiete (Methodik als Teildisziplin der Didaktik), aber auch durch eine inhaltlich beträchtlich erweiterte Fragenpalette. Diese Fragenpalette soll zur Strukturierung der folgenden Erläuterungen dienen und eine umfassende Darstellung der didaktischen Konzeption der Lernumgebung PTO gewährleisten.

1.2.3 Didaktische Fragestellungen

Wer? Damit wird die Frage nach dem Zielpublikum gestellt. PTO versteht sich als Lernangebot für Personen mit akademischem oder quasi-akademischem Bildungshintergrund, also vor allem für Studierende der Psychologie oder Medizin, aber auch für Absolventen und Absolventinnen von Fachhochschulen für therapeutische und soziale Berufe. PTO richtet sich an Lernende, die in ihrem derzeitigen oder zukünftigen Berufsleben mit Patienten konfrontiert werden, die an psychischen Störungen leiden. Dazu können beispielsweise Pflegefachleute in der Psychiatrie, Sozialpädagogen und -pädagoginnen oder Ergotherapeuten und -therapeutinnen zählen. In den meisten Fällen werden es Personen in Aus- oder Weiterbildung in institutionellem Rahmen sein. Der Zugang zu PTO ist aber

auch für Einzelpersonen möglich. Aus rechtlichen Gründen muss sich allerdings jeder Benutzer registrieren, d. h. die Lizenzgeberin muss jeden Benutzer identifizieren können. PTO existierte bei Drucklegung dieser Arbeit nur in deutscher Version. Prinzipiell könnten auch Personen mit anderer als deutscher Muttersprache zum Zielpublikum von PTO gehören, die Herstellung von anderen Sprachversionen hätte aber das von Bund und Kanton bewilligte Budget überstiegen.

Was? Der Inhaltsbereich von PTO wurde durch das Bedürfnis definiert, einen Teil des Curriculums für Klinische Psychologie und Psychotherapie durch ein computergestütztes Lernangebot abzudecken. Da Psychopathologie aber eine fachübergreifende Begrifflichkeit bieten soll, eignet sich PTO ebenso für verwandte Fächer, in denen psychische Störungsbilder eine Rolle spielen. Die inhaltlichen Grenzen wurden bereits weiter oben umrissen.

Neben der Sektion, die sich ausschliesslich mit psychischen Störungen beschäftigt, bietet PTO ein Paket von Lektionen, die störungsübergreifende Themen abdecken und von führenden Fachexperten beigesteuert wurden. Zusammen geben diese beiden Abschnitte den Lernenden einen breiten, aber übersichtlich geordneten Überblick über im klinischen Alltag relevante psychische Störungen, neben einer allgemeinen Einführung ins Fach aber auch über notwendige oder kontrovers diskutierte Metathemen.

Von wem? Diese Frage muss auf zwei Ebenen beantwortet werden: Aus Sicht der meisten Benutzer und Benutzerinnen sind die Fachlehrpersonen der betreffenden Studienrichtungen die Anbieter von PTO. Es liegt in der Verantwortung der Lehrpersonen, die PTO als Teil ihres Curriculums nutzen, die Lernumgebung sinnvoll in den Kanon von Veranstaltungen und Lernangeboten ihres Fachbereichs zu integrieren und damit ihren Unterricht zu gestalten. Das PTO-Team bietet diesbezüglich eine individuelle Beratung für interessierte Lehrpersonen an.

Aus rechtlicher Perspektive jedoch ist die Universität Zürich die Lizenzgeberin. PTO wurde am Psychologischen Institut der Universität Zürich entwickelt und wird mit Unterstützung der dortigen Informatikdienste und Kompetenzzentren betrieben. Interessierte Bildungsinstitutionen, aber auch Privatpersonen können Lizenzen zur Benutzung von PTO erwerben.

Wann? Es ist ein vielgepriesener Vorteil virtueller und insbesondere internetgestützter Lernangebote, dass die Lernenden sie zeitlich sehr flexibel und nach ihren Bedürfnissen nutzen können. Prinzipiell ist der Zugang zu PTO für berechnigte Personen jederzeit gewährleistet, und die Inhalte sind innerhalb des Passwortschutzes frei anwählbar. Den Lernenden wird allerdings empfohlen, sie in einer bestimmten Reihenfolge zu bearbeiten, da manche Inhalte auf Wissen aus anderen Lektionen aufbauen. Der Besuch beliebiger Sei-

ten in PTO ist aber an keine Bedingung geknüpft, wer über die Zugangsdaten verfügt, kann in PTO navigieren wohin und wann er oder sie will.

Mit wem? Theoretisch kann PTO im Alleingang bearbeitet werden. Allerdings wurde die Lernumgebung nicht als Stand-alone-Curriculum konzipiert, sondern als Teil eines umfassenderen Angebots in den Fachbereichen, die sich der Psychopathologie bedienen. Damit adressiert PTO in der Regel eher Gruppen von Auszubildenden, von Klassen- oder Seminargrösse bis hin zum Publikum einer Vorlesung mit über 200 Teilnehmern und wiederholten Veranstaltungen über mehrere Wochen. Mit diesen Voraussetzungen bietet es sich an, stoffbezogene Interaktionen und kooperative Lernprozesse unter den Lernenden zu initialisieren. Da es aber in der Verantwortung der Lehrpersonen liegt, wie PTO im Rahmen der Veranstaltungen eingesetzt wird, wird es diesbezüglich verschiedenste Szenarien geben. Das schwierigste Setting dürfte die Vorlesung sein, da hier die Betreuung der Studierenden aufgrund der schieren Teilnehmerzahl erhöhte Anforderungen stellt. Deshalb wurden drei prototypische Übungen konzipiert, die den Lehrpersonen Möglichkeiten aufzeigen sollen, wie PTO in Lerngruppen von je vier Lernenden eingesetzt werden kann. Die Übungen beziehen sich auf zentralen Stoff aus dem Lektionsangebot von PTO und sind bis ins Detail ausgearbeitet. Die Übungskonzepte können aber nur eine Ausgangsbasis darstellen. Jede Lehrperson wird sich, ausgehend von ihrer eigenen Prioritätensetzung, eigene Konzepte zurechtlegen müssen, wie sie die Kooperation unter den Lernenden anregen und betreuen will.

Wo? Die örtliche Unabhängigkeit beim Lernen mit PTO ist relativ. Prinzipiell steht PTO an jedem Rechner mit Internetzugang zur Verfügung, sofern die Benutzer zugriffsberechtigt sind. Da die Betreiberin die Benutzer kennen muss, existiert keine käufliche Offline-Version. Zudem müssen auf dem Rechner des Benutzers einige Programme installiert sein, die die Nutzung der herunterladbaren Dokumente und der interaktiven Möglichkeiten sowie die Darstellung der Videoinhalte ermöglichen. In der Regel aber verfügt der grösste Teil der Lernenden über Computer mit entsprechender Ausstattung, so dass lediglich die Verfügbarkeit eines Internetzugangs die Nutzer örtlich einschränkt. Mit den immer zahlreicheren WLAN-tauglichen Geräten ist die Benutzung zumindest im näheren Umfeld von Institutionen, die WLAN anbieten, derzeit schon vielerorts uneingeschränkt möglich (sofern man über die Zugangsberechtigung zum jeweiligen Netz verfügt), und auch zu Hause sind die meisten Nutzer mit einem Internetzugang ausgerüstet. PTO ist aufgrund der Grösse einiger Dateien mit einem Breitband-Internetanschluss am bequemsten zu benutzen, es wurde aber Wert darauf gelegt, dass auch Lernende mit älterem Modem die meisten Quellen noch mit akzeptablen Ladezeiten nutzen können.

Wie? Wie oben erwähnt, liegt die Entscheidung über die Einbettung von PTO ins Gesamtcurriculum in der Verantwortung der Lehrperson. Konzipiert wurde PTO mit der Vorstellung, dass (einzelne) Veranstaltungen, in denen es um die Vermittlung der Erscheinungsbilder psychischer Störungen geht, teilweise oder vollständig durch Selbststudium ersetzt werden können. Dabei wird explizit davon abgeraten, eine ganze Vorlesungsreihe durch PTO zu ersetzen. Der Glaube, E-Learning sei ein didaktischer Selbstläufer und seine Implementierung mache persönliche Betreuung und die Präsenz einer Lehrperson hinfällig, ist irrig. Wer seine Lernenden mit einem virtuellen Lernangebot alleine lässt, ohne sie in irgendeiner Form individuell anzusprechen oder den Kontakt zwischen sich und den Lernenden sowie untereinander anzuregen, wird sich rasch mit schwindenden Teilnehmerzahlen konfrontiert sehen.

Der empfohlene Einsatz von PTO sieht hingegen so aus, dass einige Veranstaltungen durch Heimstudiumsphasen mit PTO ersetzt werden, in denen die Lernenden sich den im Voraus definierten oder vereinbarten Stoff bei Sitzungen am Computer aneignen. Periodisch treffen sich Lehrperson und Lernende zu geleiteten Präsenzveranstaltungen, in denen Fragen aus der vorhergehenden Heimstudiumsphase geklärt und begleitende Themen behandelt werden, die in Zusammenhang mit PTO stehen. In der Regel wird die Selbststudiumsphase die Vorbereitung sein für die Präsenzveranstaltung, d. h. die Lernenden haben die Aufgabe, sich mit PTO bestimmte Inhalte anzueignen, die dann in der weiterführenden Präsenzveranstaltung vorausgesetzt werden können. Die Lehrpersonen delegieren also die Vermittlung des Basiswissens an PTO und können sich in der Präsenzveranstaltung auf Elaborationen, komplexe Fallbeispiele, Anwendungen oder weiterführende Themen konzentrieren. Voraussetzung ist, dass Fragen aus dem Selbststudium vorgängig geklärt werden. Um solche zentral zu sammeln, bietet PTO ein Forum an. Dieses wird durch die Grundfunktionalität von OLAT, einem technischen Rahmenwerk für E-Learning an der Universität Zürich, in das auch PTO eingebettet ist, zur Verfügung gestellt.

Normalerweise werden sich die Lernenden den Stoff in individuellen Heimsitzungen aneignen. Wie oben beschrieben, können die Lehrpersonen aber auch Übungsszenarien entwerfen, in denen die Lernenden gruppenweise zu kollaborativen Lernphasen angeleitet werden. Hier ergibt sich ein Überschneidungsbereich mit der Frage "Mit wem?".

Wie aber lernt man mit PTO in Phasen des Heimstudiums? Fälle, in denen die Lehrperson auch in diesen Phasen internetgestützte Betreuung anbietet, dürften aus personellen und finanziellen Gründen eher die Ausnahme darstellen. Die Lernenden werden also auch beim empfohlenen Szenario vorübergehend auf sich allein gestellt sein. Dies bedingt ein durchdachtes Instruktionsdesign, denn es gilt zu verhindern, dass sich die Lernenden im Stoff verlieren oder sich Misskonzepte aneignen. PTO begegnet diesen Anforderungen auf zwei Ebenen:

- Die Gliederung der störungsbezogenen Inhalte folgt dem Klassifikationssystem der ICD-10, das stark hierarchisch aufgebaut und deshalb gut über eine Webnavigation abbildbar ist. Die Lektionsstrukturen sind einfach und grösstenteils invariant, wodurch sich die Benutzer rasch zurechtfinden und den Überblick und die Orientierung behalten.
- Die zweite Massnahme steht auf kognitionspsychologischer Grundlage. Absicht ist, den Lernenden zuerst eine stabile Grundlage zu vermitteln, in die sie dann im weiteren Verlauf des Lernens mit PTO ergänzende Inhalte einbauen können, ohne dass kognitive Dissonanz auftritt. Dazu wurden die störungsbezogenen Lektionen in zwei Pakete eingeteilt. Die Basis stellen 20 klinisch relevante Störungen dar, und eine Wissensdiagnostik prüft nach der Erarbeitung dieser Störungen auf individueller Ebene, ob die Lernenden im Gedächtnis eine korrekte Überblicksstruktur der bearbeiteten Störungen aufgebaut haben. Falls die Wissensdiagnostik auf eine fehlerhafte Struktur stösst (gemessen an einer Expertenstruktur), so generiert PTO automatisiert die geeigneten Übungen und Wiederholungsempfehlungen, um die Struktur zu korrigieren (Egli, 2008).

Womit? Diese Frage ist nur auf den ersten Blick rasch beantwortet. PTO wird natürlich primär mit dem Computer bearbeitet. Allerdings bieten viele Lektionen Übungen an in Form von Fallbeispielen zum Herunterladen. Diese Fallbeschreibungen werden von den Lernenden auf ihre Symptomatik hin analysiert. Für geübte Benutzer stehen diese Beschreibungen als Word-Datei zur Verfügung, in die man seine Beobachtung direkt am Computer einträgt. Wer es traditioneller mag, druckt ein entsprechendes PDF-Dokument aus. Das Beispiel und seine Bearbeitung bleiben dann physisch verfügbar.

Referenzen im Fliesstext sind als Hyperlinks zu einem Literaturverzeichnis gestaltet. Jede Lektion verfügt über ein solches Verzeichnis, in dem sämtliche Quellen, die der entsprechenden Lektion zugrunde liegen, aufgeführt sind. PTO-Lektionen stellen eine für das Lesen am Bildschirm aufbereitete Kombination der ergiebigsten Fachlehrbuch-Texte dar, die zu einem Störungsbild gefunden wurden. Natürlich behandeln diese Lehrbücher in der Regel weit mehr, als in PTO dargestellt wird, weil sie sich meist nicht nur mit der Phänomenologie beschäftigen. PTO gibt zwar stets diese weiter führenden Quellen an, wie auch Links zu Internetseiten mit interessanten Informationen mit Bezug zum Thema. PTO weist die Lernenden aber immer wieder darauf hin, sich bei der Nutzung von solchem Material klare Erkenntnisziele zu setzen, sich also z. B. bei einem Besuch auf einer externen Internetseite nicht zu unstrukturiertem Surfverhalten verleiten zu lassen.

Wozu? Wer sich mit psychischen Störungen befassen will, muss wissen, wie sie definiert sind. Dementsprechend enthalten Lehrbücher zum Thema - sei es in Medizin, Psycholo-

gie oder anderen Fachbereichen - jeweils eine kurze Beschreibung der Symptomatik und der restlichen Kriterien, die eine Störung definieren. In Vorlesungen werden diese Inhalte meist ebenfalls vermittelt, oft in einer einsemestrigen Veranstaltungsreihe. Für die Dozierenden ist diese Lösung oft unbefriedigend. In regelmässigen Abständen muss derselbe basale Stoff angeboten werden. Oft gelingt es einer Lehrperson, mit eigenen Erfahrungsberichten und Fallbeispielen, die Vermittlung dieses an sich relativ trockenen Stoffes attraktiv zu gestalten, allerdings mit dem Nebeneffekt, dass längst nicht alle klinisch relevanten Störungen in der verfügbaren Zeit besprochen werden können. Andere Dozierende konzentrieren sich gleich auf die für sie häufig interessanteren Entstehungstheorien oder Behandlungsansätze und überlassen es den Lernenden, sich das Symptomwissen selbst anzueignen. Tatsächlich erscheinen Präsenzveranstaltungen eher als E-Learning-Module geeignet zur Vermittlung von Behandlungsmethoden, Fragen der Berufsethik oder diagnostischen Fähigkeiten. Das grosse Merkmalswissen jedoch, das in der Diagnostik unabdingbar ist, eignet sich gut zur Aneignung im Selbststudium, sofern dieses sinnvoll strukturiert wird. Ein Selbststudium birgt aber gewisse Probleme und Gefahren: Die Lernenden haben selten vor einer Zwischen- oder Abschlussprüfung Gelegenheit, die Adäquatheit ihres Wissens zu testen. Wollen sie den Stoff praxisnah und spannend erleben, kommen sie in der Regel nicht ohne Zusatzliteratur aus, und es bleibt dann ihnen selbst überlassen, das Gelesene zu strukturieren und zu gewichten. Hier ist E-Learning eine sinnvolle Lösung, denn ein entsprechend gestaltetes Lernangebot kann alles in einem Paket liefern: gut strukturierten Stoff mit klarer Navigation, interaktive Lernhilfen, Fallbeispiele in verschiedenen Formaten, eine breite, automatisierte Wissensdiagnostik, Übungen mit Kontrollmöglichkeit usw.

1.3 Lerninhalte

PTO gliedert sich grundsätzlich in zwei Bereiche: zum einen in den Bereich Lerninhalte, welcher Wissen zu den Erscheinungsbildern psychischer Störungen vermittelt. Zum anderen in den Bereich Wissensdiagnostik, welcher das Adaptive Tutorielle Systeme enthält. Die Struktur des Bereichs Lerninhalte orientiert sich am international anerkannten und in der Praxis stark verbreiteten Klassifikationsschema ICD-10⁴. Im Abschnitt F behandelt dieses psychische sowie Verhaltensstörungen, die in insgesamt 10 grössere Gruppen eingeteilt werden (F0 – F9). Die Gruppe Intelligenzminderung (F7) und die Gruppe Entwicklungsstörungen (F8) werden in PTO nicht thematisiert. Von der Gruppe Verhaltens- und emotionale Störungen mit Beginn in Kindheit und Jugend (F9) wird nur ein Störungsbild behandelt, bei dem neuere Erkenntnisse nahe legen, dass es durchaus Ver-

⁴ Das Akronym ICD steht für "International Classification of Diseases". Herausgeber des Klassifikationsschemas ist die Weltgesundheitsorganisation WHO. Die aktuelle Version ist die zehnte.

läufe bis ins Erwachsenenalter geben kann. Somit konzentriert sich PTO im Wesentlichen auf psychische Störungen des Erwachsenenalters.

Die in PTO behandelten Störungsgruppen werden den Lernenden auf der obersten Ebene zur Auswahl präsentiert (links in Abb. 1). Auf gleicher Ebene bietet der Bereich Spotlight (rechts in Abb. 1) ausgewählte Themen der Psychopathologie. Neben einer Einführung in den Fachbereich werden spezifische, relevante Problempunkte der Psychopathologie diskutiert. Hier erfolgt auch eine Thematisierung therapeutischer Aspekte. Diese werden ansonsten in PTO nicht berücksichtigt. Dies gilt grösstenteils auch für die verursachenden Faktoren (Ätiologie), welche nur dann aufgeführt werden, wenn sie Einfluss auf die Klassifikation einer Störung haben.

Der Ausdruck "Störungsgruppe" wird in PTO durchgängig zur Bezeichnung der obersten, unschärfsten Hierarchie-Ebene Fx verwendet. Dies, da die ICD-10 den Ausdruck "Kategorie" unspezifisch für alle Unterteilungsstufen gebraucht, was bisweilen zu Verwirrung führt. Deshalb wird in PTO konsequent eine spezifische Nomenklatur verwendet, die den Lernenden in der Kurshilfe kommuniziert wird. In dieser Nomenklatur enthält eine Störungsgruppe der Ebene Fx verschiedene Störungskategorien der Ebene Fxx, welche wiederum über mehrere spezifische *Störungsbilder* verfügen können. Letztere werden – mit einigen Ausnahmen – meist durch eine oder zwei weitere Stellen differenziert (Fxx.x). So ist beispielsweise die „schizoaffektive Störung, gegenwärtig manisch“ (F25.0) ein Störungsbild der Kategorie „Schizoaffektive Störungen“ (F25), welche wiederum der Gruppe F2 „Schizophrenie, schizotype und wahnhaftige Störungen“ untergeordnet ist. Die Navigation innerhalb von PTO widerspiegelt damit die hierarchische Struktur der ICD-10 (von der Gruppe zur Kategorie zum spezifischen Störungsbild). Zur leichteren und schnelleren Orientierung sind die F-Gruppen zudem farblich codiert.



Abb. 1: Übersichtsseite mit den F-Gruppen sowie dem Bereich Spotlight mit ausgewählten Themen der Psychopathologie.

Klicken die Lerner in der Gruppenübersicht (Abb. 1) auf eine Störungsgruppe (z. B. die Gruppe Affektive Störungen F3), navigieren sie auf die Seite Kategorienauswahl (Abb. 2). Hier werden ihnen die verschiedenen Störungskategorien (Ebene Fxx) der entsprechenden Gruppe präsentiert. Allerdings werden in PTO nicht alle in den Klassifikationssystemen erfassten Störungsbilder und Störungskategorien dargestellt. Merkmale wie Verbreitung in der Bevölkerung, klinische Relevanz oder Genauigkeit der Störungsdefinition dienen als Entscheidungskriterium bei der Aufnahme in PTO.

Affektive Störungen

Wählen Sie aus der folgenden Liste die Kategorie der Störungsbilder, mit der Sie sich beschäftigen möchten.

F30 **Manische Episode**

F31& F33 **Bipolare & rezidiv. depressive Störungen**

F32 **Depressive Episode**

F34 **Anhaltende affektive Störungen**

Abb. 2: Kategorienauswahl.

Auf Kategorie-Ebene gelangen die Lerner zuerst auf eine Übersicht über die Lerninhalte der betreffenden Kategorie (Abb. 3). Die Benutzer können hier wählen, ob sie sich mit Inhalten der Kategorie-Einführung (links) oder mit spezifischen Störungsbildern der Kategorie (rechts) auseinander setzen möchten.

| Elemente der Kategorie-Einführung | Spezifische Störungsbilder |
|---|--|
| <ol style="list-style-type: none"> 1. Einstieg 2. Lernziele für die Einführung 3. Allgemeines zur Kategorie 4. Allg. ICD-10 Kriterien 5. Verlauf 6. Unterschiede ICD-10/DSM-IV 7. Übung 8. Quellen (inkl. Lernziel-Check) | <ul style="list-style-type: none"> • Manie mit psychotischen Symptomen C1 • Hypomanie C2 |

Abb. 3: Übersicht über die Lerninhalte einer bestimmten Kategorie: Kategorieeinführung sowie Informationen zu spezifischen Störungsbildern.

Jede Kategorieeinführung ist nahezu identisch strukturiert und enthält die in der Abbildung aufgeführten Elemente. Ausnahmen bilden die Seiten "Allgemeine ICD-10 Kriterien" und "Übung" in Abb. 3.

Allg. ICD-10 Kriterien. Wo vorhanden, werden die generellen Kategorie-Kriterien der ICD-10 aufgeführt. Dies ist dort der Fall, wo sich spezifische Störungsbilder eher durch die erforderliche Symptomanzahl, spezielle Verläufe oder eine komorbide Symptomatik auszeichnen, sonst aber wesentliche gemeinsame Merkmale haben. Wo die ICD-10 diese allgemeinen Kriterien bereits auf Kategorie-Ebene nennt (etwa bei F20 Schizophrenie), berücksichtigt PTO dies in Form dieser speziellen Seite in der Kategorie-Einführung.

Übung. Übungen finden ihren Platz normalerweise in den störungsspezifischen Lektionen. Unterscheiden sich die Störungsbilder einer Kategorie aber nur durch die minimale Anzahl der für eine Diagnose erforderlichen Symptome, so macht es wenig Sinn, in jedem einzelnen Störungsbild wieder ein neues Beispiel derselben Symptomatik zu zeigen. In solchen Fällen wurden die Übungen in die Kategorie-Einführung ausgelagert. In den Übungen setzen sich die Lernenden praxisorientiert anhand eines oder mehrerer konkreter Beispiele mit den Symptomen und Charakteristiken der Störungen auseinander.

Abgesehen von Übungen und Kriterienseiten sind Kategorie-Einführungen stets gleich sequenziert:

Einstieg. Zu Beginn werden die Lerner in die Thematik der entsprechenden Kategorie eingeführt. Dies kann anhand eines interessanten oder bewegenden Fallbeispiels oder auch einer Selbstreflexionsübung geschehen, mit einem Video oder einem Interviewausschnitt. Der Einstieg hat vor allem motivierenden Zweck und soll den kognitiven Rahmen vorbereiten.

Lernziele. Hier werden die Lernziele der Lektion definiert. Sie formulieren möglichst konkret und messbar, zu welcher Leistung die Lernenden am Schluss der Lektion imstande sein und welche Schritte sie bearbeitet haben sollten.

Allgemeines zur Kategorie. Danach erhalten die Lernenden grundsätzliche Informationen über die wesentlichen Merkmale der Kategorie. Hier wird die Symptomatik der Kategorie in leicht lesbarer Form dargestellt, meist anders strukturiert als in der ICD-10 und ergänzt durch weitere Inhalte zur Symptomatik.

Verlauf. Die meisten Kategorien von Störungsbildern verfügen über einen spezifischen Störungsverlauf (etwa die rezidivierenden affektiven Episoden der Kategorien F31 und F33).

ICD-10/DSM-IV. Neben der ICD-10 ist das DSM-IV⁵ eines der gebräuchlichsten Klassifikationssysteme, v. a. für die psychologisch orientierte Forschung. Viele Fachleute mit psychologischem Hintergrund ziehen der besseren empirischen Grundlage und der ge-

⁵ Beim DSM-IV handelt es sich um das diagnostische und statistische Manual psychischer Störungen der American Psychiatric Association APA in der vierten Version.

neueren Störungsdefinitionen wegen das DSM der ICD vor. Um diesem Umstand Rechnung zu tragen, werden die Unterschiede in der Beurteilung der Störungsbilder, die zwischen ICD-10 und DSM-IV bestehen, an entsprechender Stelle erläutert.

Quellen. Unter diesem Punkt wird diejenige Fachliteratur aufgeführt, welche der Zusammenstellung der Inhalte zu Grunde lag. Hier finden sich auch Hinweise zu allfälligen Videobeispielen sowie Verweise zu interessanten Internetseiten, wo das Gelernte vertieft werden kann.

In die Quellenseite eingebettet ist auch der sogenannte *Lernziel-Check*. Damit können die Lernziele noch einmal aufgerufen werden. In einem Pop-up-Fenster ist jedes Lernziel mit einem Klick-Icon versehen. Kann man die im Lernziel definierte Leistung nicht erbringen, so kann man sich per Klick auf das zugehörige Icon auf die Seite mit der benötigten Information führen lassen.

1.4 Lernerfreundliche Ausgestaltung

Die Inhalte von PTO sind in den verschiedensten Medienformaten (Text, Grafik, Video, Audio) gehalten. Da es sich bei PTO um eine webbasierte Lernumgebung handelt, bedingt deren Verwendung die Verfügbarkeit eines Computers mit Internet-Browser und spezifische Software-Voraussetzungen. Die Lernumgebung ist so konzipiert, dass die nötigen Zusatzprogramme zum Darstellen der verschiedenen technischen Formate frei erhältlich sind und deren Installation keinerlei Fachkenntnisse erfordert. Da einige der Lernmaterialien relativ datenintensiv sind, empfiehlt sich für eine optimale Nutzung eine schnelle Standleitung. PTO nutzt das Open-Source Learning-Management-System (LMS) OLAT⁶. Dabei handelt es sich um eine netzbasierte Lernplattform, welche von einem Server aus Lerninhalte und Lernaktivitäten anbietet und so die hohe Adaptierbarkeit und Flexibilität bei der Gestaltung von PTO erst ermöglicht.

Die Inhalte zu den spezifischen Störungsbildern bilden die eigentlichen Lektionen von PTO. Diese sind analog zur Kategorieebene strukturiert. Die konsistente Verwendung desselben Aufbaus reduziert die Komplexität und hilft den Lernern bei der Orientierung. Die Formulierung der Inhalte erfolgt zielgruppennah und den Anforderungen einer virtuellen Lernumgebung angepasst. Dies zeigt sich vor allem in der Verwendung kurzer, prägnanter Sätze, einfacher Satzstrukturen und dem Verzicht auf Fremdwörter, wo diese nicht notwendig sind. Die reinen Text- und Bildinhalte von PTO stehen in Form von PDF's⁷ auch zum Herunterladen bereit (inklusive aktivierter Kommentar- und Markierfunktion im Adobe Reader). Dadurch soll den Lernern die Möglichkeit gegeben werden,

⁶ <http://www.olat.uzh.ch>

⁷ Das Akronym PDF steht für „Portable Document Format“. Dabei handelt es sich um ein proprietäres plattformübergreifendes Dateiformat für Dokumente, das von der Firma Adobe Systems entwickelt wurde.

sich bei Bedarf bestimmte Inhalte auch ausdrucken zu können. Zu jeder Übung steht ebenfalls ein Dokument zum Herunterladen und Ausdrucken bereit. In diesem können die Lerner die Beobachtungen, namentlich die Symptome und Kriterien, die sie erkennen, niederschreiben. Zudem finden sich zu den meisten Übungen Musterlösungen zur Selbstkontrolle. Die Inhalte in PTO werden meist in Kombination mit Abbildungen präsentiert. Diese Doppelkodierung der Inhalte in Wort und Bild lockert die Bearbeitung auf und führt zu einer stärkeren Verankerung des Gelernten. Die relevantesten Störungsbilder sind weiter mit Fallbeispielen in verschiedenen Medienformaten versehen, um so den Lernern handlungsorientiert die Symptome anhand konkreter, lebendiger Beispiele näher zu bringen. In PTO kommen an verschiedenen Stellen interaktive Grafiken zum Einsatz. Die lernrelevanten Inhalte können dabei jeweils zur Selbstabfrage ausgeblendet werden. Überdies können mit einer Speed-Reading-Funktion im Fliesstext die relevanten Textstellen hinterleuchtet werden. Dies ermöglicht den Lernern einen schnellen Überblick über die Kernaussagen eines Abschnitts und erleichtert die Repetition. Ein weiteres im Rahmen von PTO entwickeltes Element ist das Pop-up-Glossar. Dabei werden Definitionen von Fachbegriffen durch Positionierung des Mauszeigers auf dem entsprechenden Wort direkt angezeigt. Dies macht ein Nachschlagen der Begriffe überflüssig, vermeidet unnötige Unterbrechungen und führt zu effizienterem Lernen.

Zu den Inhalten der einzelnen Lektionen sowie zum Spotlight werden Selbsttests angeboten. Diese sind teils in Form von Single- und Multiple-Choice-Fragen, teils aber auch in Form von Lückentextaufgaben gestellt. Die Selbsttests geben – neben der relationalen Diagnostik des strukturellen Wissens im Bereich Wissensdiagnostik – den Lernern auf deklarativer Ebene die Möglichkeit, jederzeit ihr Wissen zu den Inhalten von PTO zu prüfen. PTO verfügt weiter über ein Forum, welches einerseits dazu dient, sich bei spezifischen technischen und inhaltlichen Fragen an die Betreiber zu wenden, welches aber auch für fachspezifischen Austausch zwischen den Lernenden untereinander und zwischen Lernenden und Dozierenden respektive deren Mitarbeitern genutzt werden kann.

1.5 Der Lernprozess in PTO

Der Lernprozess innerhalb von PTO umfasst zwei störungsbezogene Curriculumphasen, die aufeinander aufbauen und welche die Lernenden idealiter sequenziell bearbeiten. Das erste Curriculum (C1) enthält eine repräsentative Auswahl der 20 relevantesten psychopathologischen Störungsbilder, die das gesamte phänomenologische Spektrum an Störungen abdecken. Eine Vorauswahl von Störungsbildern erfolgte dabei nach den Kriterien der klinischen Prävalenz und Relevanz. Zudem sollten die gewählten Störungsbilder das gesamte Spektrum psychischer Störungen abdecken. Diese Vorauswahl wurde den in der Schweiz führenden Leherxperten der Psychopathologie zur erneuten Beurteilung vorge-

legt. Auf diese Weise wurde schliesslich auf 20 Störungsbilder für C1 reduziert (vgl. Egli, 2008).

Bei der Bearbeitung von C1 geht es um den Aufbau einer stabilen und korrekten Grundstruktur, in welche später im Sinne assimilativen Lernens die weiteren Störungsbilder des zweiten Curriculum (C2) eingefügt werden können. Ein solches Vorgehen reduziert den Grad an Akkomodation, welcher bei der Anknüpfung von neuem an bestehendes Wissen anfällt, auf ein Minimum (Ausubel, 1980/81, 2000; Piaget, 1976).

C2 bietet einen Grobüberblick über 32 weitere Störungsbilder. Einige davon gehören bereits zu in C1 eingeführten Kategorien, während andere Teil neu eingeführter Kategorien sind. Die Inhalte von C2 sind im Vergleich zu denen von C1 knapper gehalten. Das Ziel bei der Bearbeitung von C2 besteht primär darin, einen breiteren Überblick über die verschiedenen Störungsbilder zu entwickeln und damit eine fundierte Basis für eine Ausbildung im Bereich Psychopathologie zu erhalten.

1.6 Adaptives Lernen mit PTO

Die Wissensdiagnostik mit dem Adaptiven Tutoriellen System (ATS) in PTO bietet den Lernenden die Möglichkeit, ihr Wissen zu den psychopathologischen Störungsbildern mit einem innovativen Instrument zur Wissensdiagnostik periodisch zu erheben und mit Expertenwissen zu vergleichen. Dadurch zeigen sich den Lernenden Mängel im eigenen Wissen. Zusätzlich erhalten sie konkrete Lernempfehlungen sowie spezifische Übungen zur Behebung dieser Wissensmängel. Diese Individualisierung im Lernprozess führt zu Effizienzgewinn und höherer Effektivität und fördert Motivation und Aufmerksamkeit. Im Folgenden soll erläutert werden, wie das Adaptive Tutorielle System von PTO konzipiert und aufgebaut ist, auf welchen Grundlagen und technischen Verfahren es basiert und welche Möglichkeiten sich daraus für die Lernenden ergeben.

1.6.1 Assimilation und Akkomodation beim Lernen

Der Aufbau von PTO fusst auf kognitionspsychologischen Theorien zu Assimilation und Akkomodation beim Lernen. Aus dieser Sicht repräsentiert sich das Wissen eines Menschen in kognitiven Strukturen. Die einzelnen Wissens Elemente stehen nicht isoliert für sich, sondern sind in Strukturen organisiert und geordnet. Neues Wissen wird somit in bereits bestehende Wissensschemata eingefügt. Dieser Prozess der Eingliederung von Information und Reizen aus der Umwelt in vorhandene Gefüge wird von Piaget (1976) mit Assimilation bezeichnet. Das Hinzukommen neuen Wissens hat jedoch Konsequenzen für die gesamte Wissensstruktur und führt zu Anpassung und Modifikation derselben. Diese Änderungen im Strukturgefüge auf Grund eines neu gelernten Elements nennt Piaget Akkomodation. Akkomodation und Assimilation werden von ihm als

Teilprozesse bei der Bildung eines kognitiven Gleichgewichts im Wissenserwerb verstanden. Wird diese Theorie als Grundlage für die konzeptuelle Ausarbeitung von PTO herangezogen, sollte die Wissensdiagnose innerhalb von PTO dieser Sichtweise Rechnung tragen und entsprechend ausgerichtet sein. Daher wurde diese so konzipiert, dass sie ebenfalls auf die Wissensschemata abzielt und die Wissensstruktur zu den psychopathologischen Störungsbildern erfasst. Eine bloße Erhebung des reinen Faktenwissens, beispielsweise mittels Multiple-Choice-Fragen, greift hier zu kurz. Entsprechend wurde mit der Nonmetrischen Multidimensionalen Skalierung (NMDS) auf eine Methode zurückgegriffen, welche die relationalen Beziehungen der einzelnen Wissens Elemente in Form Kognitiver Karten als geometrische Visualisierung modelliert.

1.6.2 NMDS und Kognitive Karten

Das deklarative Merkmalswissen einer Person über eine Anzahl von Objekten lässt sich mit sogenannten Kognitiven Karten darstellen (Läge, 2001, Marx & Hejj, 1989, Marx & Läge, 1995). Die Datengrundlage zur Generierung einer solchen Karte liefern paarweise Ähnlichkeitsvergleiche zwischen den einzelnen Störungsbildern. Abb. 4 zeigt die entsprechende Benutzerschnittstelle zur Erhebung der Ähnlichkeitsurteile, wie sie in PTO zur Anwendung kommt.

Anorexia nervosa

Soziale Phobien

minimale Ähnlichkeit maximale Ähnlichkeit

Abb. 4: Benutzeroberfläche von PTO zur Erhebung der Ähnlichkeitsurteile.

Die Lerner geben ihre Urteile zur Phänomenologie psychopathologischer Störungsbilder auf der Basis ihres Merkmalswissens auf einer neunstufigen Skala ab. Die globalen Ähnlichkeitseinschätzungen zwischen je zwei Objekten bildet eine Person weitestgehend

automatisiert auf der Basis der berücksichtigten Merkmale und deren Ausprägungsunterschiede (vgl. Klauer, 1989 und Smith, Shoben & Rips, 1974). Ein Ähnlichkeitsurteil bezieht auf effiziente Art und Weise das Wissen über zwei Objekte mit ein. Ein Erhebungsdurchgang bei 20 Objekten (dies entspricht 190 Urteilen) dauert erfahrungsgemäss zwischen 10 und 20 Minuten.

Mittels NMDS lassen sich diese Ähnlichkeitsdaten in einer Proximitätenmatrix als Distanzwerte verrechnen, in ein Raummodell überführen und als Kognitive Karte veranschaulichen (Borg & Groenen, 1997; Läge, 2001). Als dimensionale Strukturgefüge stellen sie die Relationen, die eine Person zwischen diesen Objekten sieht, in Form von kleineren oder größeren Distanzen dar und lassen als Gesamtbild Rückschlüsse auf die individuelle Gedächtnisrepräsentation zu. Abb. 5 zeigt die kognitive Struktur für den Wissensbereich psychopathologischer Störungsbilder eines Lernalers in einem zweidimensionalen Raum⁸.

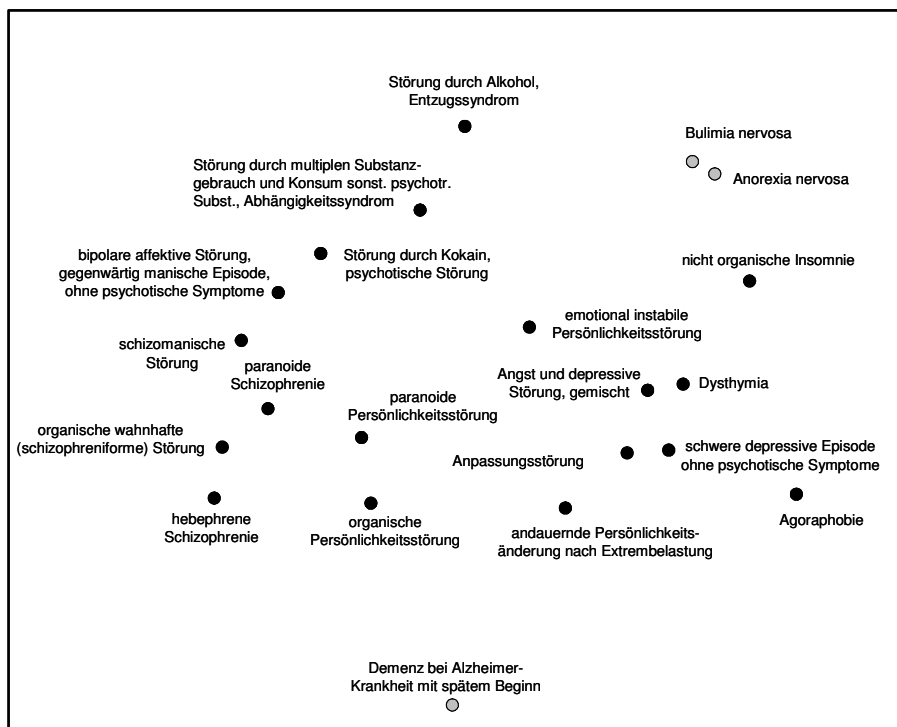


Abb. 5: Zweidimensional skalierte NMDS-Lösung. Die Punkte repräsentieren die Position der psychopathologischen Störungsbilder (graue Punkte markieren die unten im Text beschriebenen Beispiele).

⁸ Zur Erstellung der Kognitiven Karte wurde der robuste NMDS-Algorithmus RobuScal (Läge, Daub, Bosia, Jäger und Ryf, 2005) verwendet. Dieser bietet durch die Verwendung eines gewichteten Fehlermodells die Möglichkeit, auf individueller Ebene Karten zu interpretieren.

Nah beieinander positionierte Störungsbilder wurden als ähnlich bewertet, weit auseinander liegende dagegen als unähnlich. Ein Beispiel soll das Gesagte verdeutlichen: Die Anorexia nervosa (Magersucht) und die Bulimia nervosa (Ess- und Brechsucht) sind phänomenologisch sehr ähnliche Störungen und gehören innerhalb des internationalen WHO-Klassifikationsschemas ICD-10 derselben Kategorie an. Die hoch eingeschätzte Ähnlichkeit spiegelt sich in der Karte in einer nahen Positionierung wieder. Die Bulimia nervosa und die „Demenz bei Alzheimer-Krankheit mit spätem Beginn“ hingegen verfügen über wenig Gemeinsamkeiten. Die hohe Unähnlichkeitsbeurteilung äußert sich in einer weit entfernten Positionierung.

1.6.3 Expertenmodell und Prokrustes-Transformation

Die führenden Textbuchautoren der Schweiz im Bereich Psychopathologie und ihre Mitarbeiter (insgesamt 20 Personen) wurden gebeten, die in PTO behandelten Störungsbilder, mittels des oben beschriebenen Ähnlichkeitsverfahrens auf phänomenologischer Ebene einzuschätzen. Aus der Matrix, der über alle Experten gemittelten Ähnlichkeitsurteile wurde mit robuster NMDS eine zweidimensionale Kognitive Karte gerechnet. Das resultierende Expertenmodell basiert auf der gemittelten Meinung von Experten im Bereich der Psychopathologie (siehe Abb. 6).

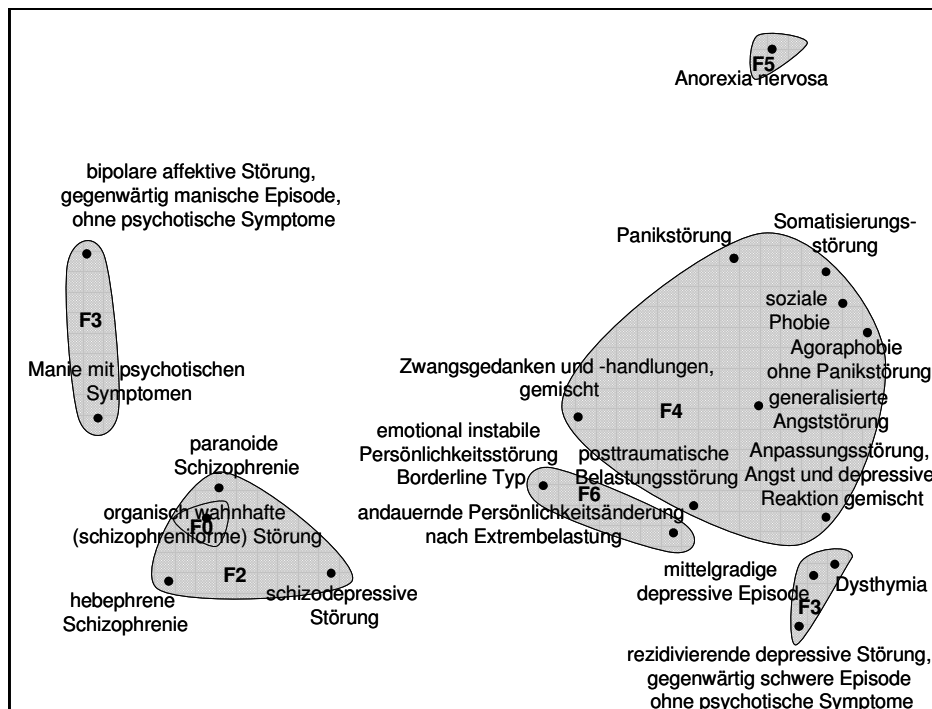


Abb. 6: Expertenmodell. Die eingezeichneten Cluster kennzeichnen Störungen, die im internationalen Klassifikationssystem ICD-10 derselben Kategorie angehören.

In dieser Kognitiven Karte sind ihre Beurteilungen der phänomenologischen Ähnlichkeit der Störungsbilder in den relationalen Proximitäten widergespiegelt. Damit repräsentiert diese Karte die Vorstellungen der Experten über die strukturelle Anordnung der Störungsbilder. Mit dem Expertenmodell steht im Rahmen von PTO eine Zielstruktur für die Lerner zur Verfügung. Es gilt demnach, die Lernerstruktur im Lernprozess der Expertenstruktur anzugleichen. Der Vergleich der aktuellen Lernerstruktur mit der Expertenstruktur zeigt Wissensdiskrepanzen und detektiert Wissensdefizite. Das Verfahren, welches einen Vergleich der Lernerkarte mit der Expertenkarte ermöglicht, ist die Prokrustestransformation (Gower & Dijksterhuis, 2004). Diese dreht, spiegelt, streckt und staucht die Lernerkarte unter Einhaltung der relationalen Objektbeziehungen solange, bis die bestmögliche Übereinstimmung mit der Expertenkarte gefunden ist (Abb. 7). Die verbleibenden Positionsunterschiede zwischen den korrespondierenden Objekten bestimmen dann das Mass der Abweichung. In der entstehenden Karte zeigt sich, welche Objekte ein Lerner gut kennt (das heisst korrekt verortet) und welche er falsch platziert. Die Abstandsinformation zwischen Ist- und Soll-Position lässt sich auch numerisch ausdrücken. Auf Kartenebene entspricht dies dem AverageLoss-Wert, während es sich auf Objektebene um den ObjectLoss-Wert handelt. Übersteigt das Abweichungsmass ObjectLoss eines Objektes in der Lernerkarte ein vordefiniertes Kriterium, interpretiert das System dieses als noch mangelhaft gewusst. Die Festlegung der Abweichung auf numerischer Ebene erlaubt zudem eine vollständig automatisierte Auswertung und damit den Einsatz des Verfahrens im Rahmen von computergestütztem Unterricht.

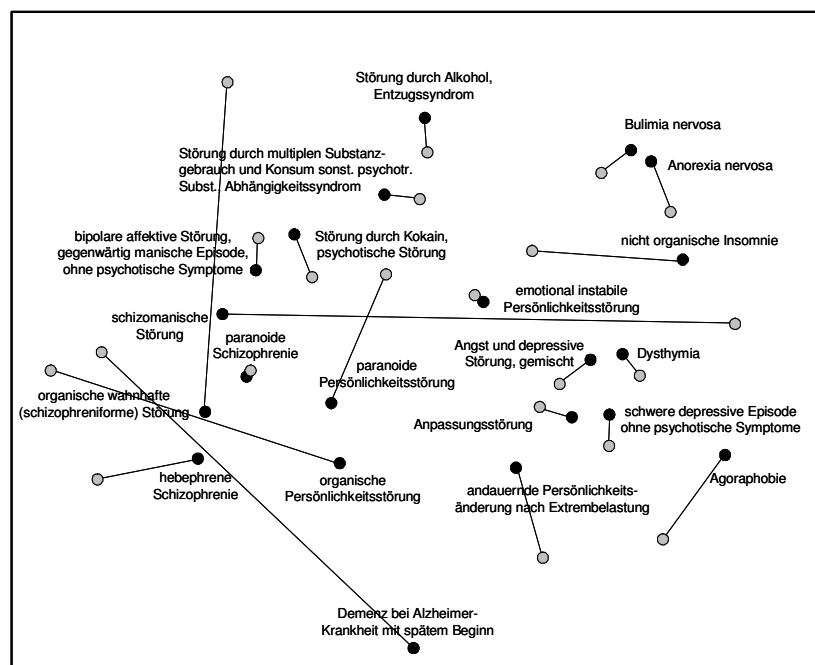
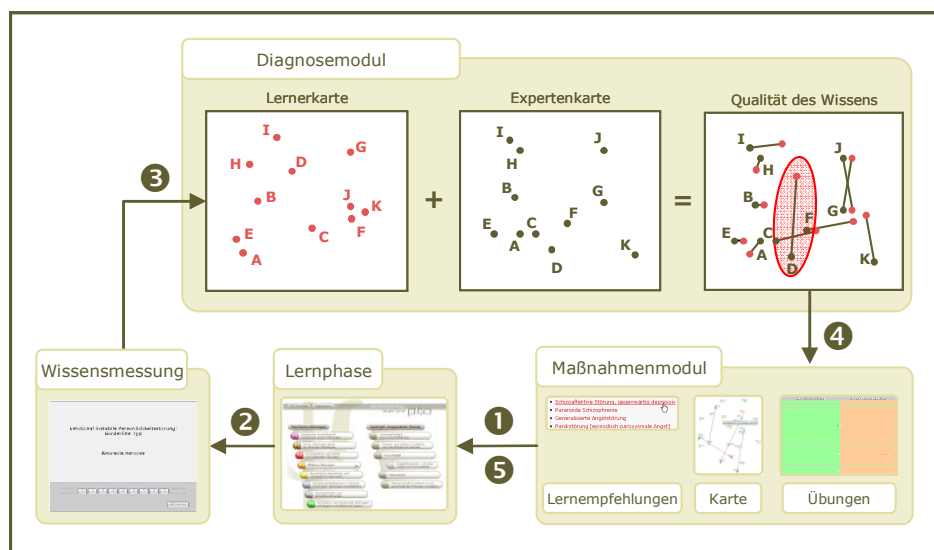


Abb. 7: Ergebnis einer Prokrustestransformation zwischen der Wissenskarte eines Lerners (graue Punkte) mit derjenigen eines Experten (schwarze Punkte).

Mit dem Expertenmodell und der Prokrustes-Transformation stehen die nötigen Mittel zur Verfügung, um ausgehend von der Wissensdiagnose eines Lerners mangelhaftes Wissen zu detektieren und dem Lerner spezifische Empfehlungen abgeben zu können.

1.6.4 Der Prozess der Wissensdiagnose und adaptive Lernempfehlungen

Der Bereich Wissensdiagnostik dient den Lernenden zur Durchführung der Wissensdiagnose. Mit dem Lernprotokoll steht ihnen zudem ein Instrument zur Steuerung und Überwachung ihres Lernprozesses zur Verfügung. Dort erhalten sie auch Zugang zu den Resultaten bereits durchgeführter Wissensdiagnosen. Möchte ein Lerner sein Wissen testen, führt er eine Wissensdiagnose durch. Wie oben beschrieben erfolgt diese ähnlichkeitsbasiert über Paarvergleiche der einzelnen Störungsbilder. Aus den resultierenden Daten wird eine Kognitive Karte generiert, welche mit einer Expertenkarte verglichen wird. Dieser Vergleich deckt Wissensmängel beim Lerner auf. Der Lerner erhält zur Schließung dieser Wissenslücken spezifische Lernempfehlungen (Streule, Egli, Oberholzer und Läge 2005). Abb. 8 veranschaulicht diesen Prozess.



- 1 / 5 Lernphase
- 2 Wissensmessung
- 3 Berechnung Wissenskarte und Detektion von Wissensmängeln
- 4 Zusammenstellung adaptiver Massnahmen

Abb. 8: Prozess der Wissensdiagnose.

Das Feedback, das die Lerner erhalten, umfasst drei Massnahmen: Wird bei der Wissensdiagnose das Lernziel (genügende Übereinstimmung der Positionen eines Störungsbildes in der individuellen Karte und der Expertenkarte) nicht erreicht, werden

den Lernenden die am weitesten abweichenden vier Störungsbilder zur nochmaligen Bearbeitung empfohlen (Häne, Streule, Egli, Oberholzer und Läge, 2006). Das mangelhafte Wissen soll durch die Repetition korrigiert werden. Zudem werden ihnen als zweite Massnahme Übungen abgegeben. Diese sind so aufgebaut, dass die Lerner auf Gemeinsamkeiten von beziehungsweise Unterschiede zwischen einem Paar von Störungsbildern hingewiesen werden, von denen genau eines in seinem Wissen relational falsch verankert ist. Dies soll dazu dienen, die strukturelle Organisation des bislang falsch verankerten Störungsbildes durch Vergleich mit einem korrekt verankerten zu verbessern. Erreichen die Lerner nach wiederholten Übungen in der Wissensdiagnose das Lernziel immer noch nicht, so wird ihnen demonstrativ mit Hilfe eines geometrischen Vergleichs ihrer eigenen Kognitiven Karte und der Expertenkarte gezeigt, wo die mangelhaft gewussten Störungsbilder innerhalb der Struktur zu liegen kommen sollten. Am Ende des Prozesses soll ja neben einem korrekten Faktenwissen auch eine korrekte relationale Verankerung der einzelnen Wissenskomponenten zueinander bestehen.

Der Nutzen, der sich durch den Einsatz dieses adaptiven Systems für die Lerner ergibt, liegt in den individuell ausgerichteten Lernempfehlungen, welche ein gezieltes, spezifisch an Wissensmängeln orientiertes Lernen ermöglichen. Dies führt zu Fokussierung auf das Wesentliche und verhindert überflüssiges Mehrfachlernen. Weiter wirkt sich ein individuelles, visuelles Wissens-Feedback positiv auf die Motivation der Lerner aus. Letztlich ermöglicht das Adaptive Tutorielle System von PTO eine nutzenmaximierende Strukturierung des Lernprozesses.

1.6.5 Wissensdiagnostik

Der Bereich Wissensdiagnostik von PTO enthält das oben vorgestellte Adaptive Tutorielle System. Die Lerner führen hier ein Lernprotokoll, können eine Diagnose ihres strukturellen Wissens zu den Störungsbildern vornehmen und erhalten individuell ausgerichtetes Feedback in Form von Lernempfehlungen und spezifischen Übungen.

1.6.5.1 Lernprotokoll

Das Lernprotokoll enthält die Störungsbilder der beiden Curricula aufgelistet (Abb. 9). Die Lerner können hier im Sinne einer Lernkontrolle festhalten, welche Störungsbilder sie bearbeitet haben und welche noch nicht. Führt ein Lerner eine Wissensdiagnose durch, wird ihm in einer separaten Spalte im Lernprotokoll zu jedem Störungsbild angegeben, zu welchem Prozentsatz er dieses korrekt verortet hat. Die Störungsbilder werden zudem geordnet nach dem Grad an Korrektheit aufgelistet. Dadurch ist dem Lerner auf einen Blick ersichtlich, bei welchen Störungsbildern er die grössten Wissensdefizite aufweist.

Wissensdiagnostik 1 - Lernprotokoll pto

Lernprotokoll
Urteilsabgabe
Instruktionen

Neuestes Lernprotokoll

Wählen Sie unten, welches **Lernprotokoll** oder **Kartenfeedback** Sie anschauen möchten (erst möglich, wenn Sie mindestens eine Urteilsabgabe durchgeführt haben).

| | | | | |
|------------------------|------------------|------------------|------------------|------------------|
| Neuestes Lernprotokoll | Lernprotokoll 4 | Lernprotokoll 3 | Lernprotokoll 2 | Lernprotokoll 1 |
| | Kartenfeedback 4 | Kartenfeedback 3 | Kartenfeedback 2 | Kartenfeedback 1 |

Hier blenden Sie eine Kurzinstruktion ein.

Ausführlichere Erläuterungen zur Nutzung des Lernprotokolls finden Sie unter **Instruktionen zum Lernprotokoll**.

Lernkontrolle

sortieren

Störungsbilder

sortieren nach: F-Code / Alphabet

Bewertung

sortieren

Angaben speichern

| | | | |
|-------------------------------------|---|----------------------------|-------|
| <input checked="" type="checkbox"/> | Bipolare affektive Störung, gegenwärtig manische Episode ohne psychotische Symptome F31.1 | Gemeinsamkeit Gegensatz | ! 0 % |
| <input type="checkbox"/> | Depressive Episode F32.0 | Gemeinsamkeit | ! |

Abb. 9: Lernprotokoll.

Die obersten vier Störungsbilder sind gleichzeitig die vier am schlechtesten gewussten und werden zur Repetition empfohlen. Zu diesen werden gleichzeitig spezifische Übungen angeboten. Überdies können aus dem Lernprotokoll jederzeit die Kartenfeedbacks der absolvierten Wissensdiagnosen zusammen mit den Lernprotokollen des entsprechenden Durchgangs aufgerufen werden. Mit dem Lernprotokoll verfügen die Lerner damit über ein Werkzeug zur Überwachung und Kontrolle ihres Lernprozesses.

1.6.5.2 Relationale Wissensdiagnose

Die eigentliche Wissensdiagnose funktioniert wie oben beschrieben ähnlichkeitsbasiert über den paarweisen Vergleich je zweier Störungsbilder (siehe Abb. 10, links). Durch diese Form der relationalen Wissensdiagnose wird das strukturelle Wissen erhoben. Als Feedback erhalten die Lerner ein Abbild ihrer kognitiven Struktur zu den erhobenen Störungsbildern (Abb. 10, Mitte). Ein Farbcode informiert sie über die Qualität der Position eines Störungsbilds innerhalb des Wissensgefüges. Zudem werden ihnen die vier am schlechtesten gewussten Störungsbilder zur Repetition empfohlen. Wird nach dreimaliger Erhebung die angestrebte Normstruktur nicht erreicht, werden die Lerner demonstrativ auf die korrekte Position dieser vier am schlechtesten gewussten Störungsbilder hingewiesen (Abb. 10, rechts).

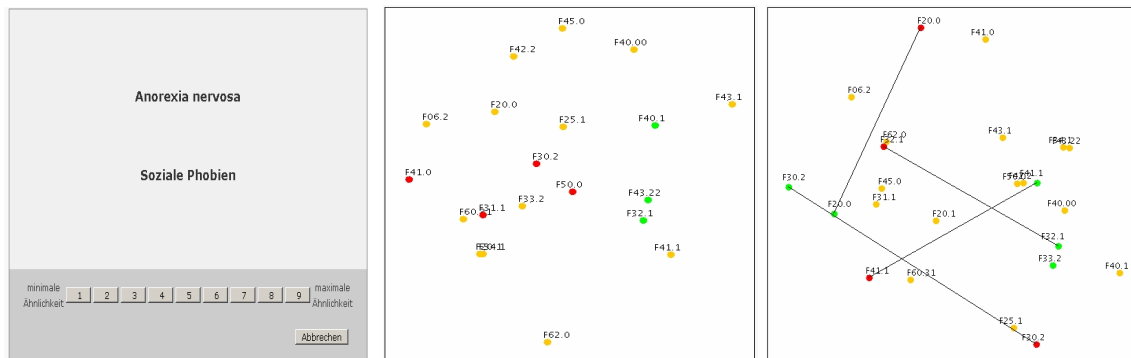


Abb. 10: Relationale Wissensdiagnose; Kartenfeedback; Kartenfeedback inklusive Hinweis der korrekten Position der vier am schlechtesten gewussten Störungsbilder (von links nach rechts).

Der Bereich Wissensdiagnostik orientiert sich am curricularen Aufbau von PTO und dient den Lernern der Kontrolle, der Steuerung und der Überwachung ihrer Lernaktivitäten im Sinne eines effizienten und nutzenmaximierten Lernprozesses. Curriculum 1 endet, wenn der akkomodative Prozess der Ausbildung einer korrekten Kognitiven Karte abgeschlossen ist. Während der Bearbeitung von Curriculum 2 findet erneut eine Wissensdiagnose statt. Diese erfolgt jedoch in einem kontinuierlichen Sinne einzeln für jedes zusätzlich gelernte Störungsbilds von C2. Dadurch soll sichergestellt werden, dass das neu gelernte Wissen richtig verarbeitet und korrekt in die bestehenden Strukturen integriert wird. Mit dem curricularen Aufbau bietet PTO einen an kognitionstheoretischen Kriterien orientierten Lernablauf. Selbstverständlich ist es den Lernenden jedoch frei gestellt, welche Inhalte von PTO sie in welcher Reihenfolge bearbeiten möchten. Auch dies ist ein Vorteil von E-Learning, dass die Lerner selbstorganisiert einen an ihren individuellen Bedürfnissen ausgerichteten Lernprozess beschreiten, worin sie der modulare Aufbau von PTO zusätzlich unterstützt.

1.7 Literaturverzeichnis

- Ausubel, D. P. (1980/81). *Psychologie des Unterrichts (Band 1 und 2)*. Weinheim: Beltz.
- Ballstaedt, S. P. (1997). *Wissensvermittlung. Die Gestaltung von Lernmaterial*. Weinheim: Beltz PVU.
- Baumann, U. & Perrez, M. (2005). *Lehrbuch Klinische Psychologie – Psychotherapie*. Bern: Hans Huber.
- Borg, I. & Groenen, P. (1997). *Modern multidimensional scaling – Theory and applications*. New York: Springer.
- Edelmann, W. (2000). *Lernpsychologie*. Weinheim: Beltz PVU.
- Egli, S. (2008). *Ein ähnlichkeitsbasiertes kognitives und klinisches Strukturmodell der psychischen Störungen*. Unveröff. Dissertation, Universität Zürich, Philosophische Fakultät.
- Gower, J. C. & Dijksterhuis, G. B. (2004). *Procrustes Problems*. Oxford: Oxford University Press.
- Häne, M., Streule, R., Egli, S., Oberholzer, R. & Läge, D. (2006). Adaptivität und deren Evaluation im E-Learning - Das Fallbeispiel "Psychopathology Taught Online" (PTO). In E. Seiler Schiedt, S. Kälin & C. Sengstad (Hrsg.), *E-Learning - alltagstaugliche Innovation? Medien in der Wissenschaft, Band 38* (S. 296-305). Münster: Waxmann.
- Jank, W. & Meyer, H. (2002). *Didaktische Modelle*. Berlin: Cornelsen.
- Klauer, K. J. (1989). Allgemeine oder bereichsspezifische Transfereffekte eines Denktrainings. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 21, 185-200.
- Läge, D. (2001). *Ähnlichkeitsbasierte Diagnostik von Sachwissen*. Habilitationsschrift an der Philosophischen Fakultät der Universität Zürich.
- Läge, D., Daub, S., Bosia, L., Jäger, C. & Ryf, S. (2005). *Die Behandlung ausreißerbehafteter Datensätze in der Nonmetrischen Multidimensionalen Skalierung - Relevanz, Problemanalyse und Lösungsvorschlag*. Forschungsberichte aus der Angewandten Kognitionspsychologie Zürich, Nr. 21. Zürich: Universität Zürich.
- Marx, W. & Heij, A. (1989). *Subjektive Strukturen*. Göttingen: Hogrefe.
- Marx, W. & Läge, D. (1995). *Der ideologische Ring*. Göttingen: Hogrefe.
- Möller, H. M., Laux, G., & Deister, A. (2005). *Psychiatrie und Psychotherapie*. Stuttgart: Thieme.
- Piaget, J. (1970). *Einführung in die genetische Erkenntnistheorie*. Frankfurt a. M.: Suhrkamp.
- Piaget, J. (1976). *Die Äquilibration der kognitiven Strukturen*. Stuttgart: Klett.
- Schermer, F. (1998). *Lernen und Gedächtnis*. Stuttgart: Kohlhammer Urban.
- Seel, N. M. (2000). *Psychologie des Lernens. Lehrbuch für Pädagogen und Psychologen*. München: Reinhardt UTB.
- Smith, E. E., Shoben, E. J. & Rips, L. J. (1974). Structure and process in semantic memory. A featural model for semantic decisions. *Psychological Review*, 81, 214-241.

Streule, R., Egli, S., Oberholzer, R. & Laege, D. (2005). Adaptive Wissensvermittlung am Beispiel der e-Learning-Umgebung "Psychopathology Taught Online (PTO)". In D. Tavangarian & K. Nölting (Hrsg.), *Auf zu neuen Ufern! E-Learning heute und morgen*. (Bd. 34, S. 47-56). Rostock: Waxmann.

2.

Theoretische Grundlagen

Abstract

Im folgenden Kapitel wird eine Annäherung an den Evaluationsbegriff unternommen. Es werden die theoretischen Grundlagen der Evaluation und der Evaluationsforschung dargestellt und in einem kurzen Abriss die historischen Gegebenheiten erläutert. Weiter werden verschiedene Arten und Typen von Evaluation vorgestellt, deren Funktion erklärt sowie die konkrete Anwendung von Methoden im Prozess der Evaluation erläutert.

In einem zweiten Teil wird spezifisch die Evaluation im Bereich von E-Learning behandelt, und zwar mit Blick auf die Evaluation der webbasierten Lernumgebung PTO. Es sollen die Grundlagen gelegt werden, anhand derer PTO systematisch und theoriegeleitet evaluiert worden ist.

2.1 Grundlagen der Evaluation

2.1.1 Einleitung

In den 70er Jahren des vergangenen Jahrhunderts hat sich erst in den USA und später auch in Europa ein Evaluationsverständnis durchgesetzt, welches sich nicht bloss auf wissenschaftliche Begleitung oder Wirkungsforschung reduzierte. Evaluation wurde neu als Mittel begriffen, welches der Verbesserung der gesellschaftlichen Praxis dient. Dies erreicht sie, indem sie Programme, Massnahmen oder Materialien systematisch, d.h. methodisch angeleitet und anhand von Gütekriterien überprüfbar beschreibt und bewertet (Beywl, 1988). Nach Stufflebeam (2007, S.4) hat Evaluation grundlegende Bedeutung: „Evaluation arguably is society’s most fundamental discipline. It is oriented to assessing and helping to improve all aspects of society. ... It permeates all areas of scholarship, production, and service and has important implications for maintaining and improving services and protecting citizens in all areas of interest to society.”

Ausgehend vom Bildungswesen fand die Evaluationsforschung Anwendung in verschiedensten Gebieten wie dem Gesundheitswesen, der Verwaltung oder der Wirtschaft. Seit einigen Jahren gehört die wissenschaftliche Evaluationsforschung denn auch als Teilgebiet der sozialwissenschaftlichen Forschung zu einem der am stärksten wachsenden akademischen Tätigkeitsfelder (Meyer & Höhns, 2002). Die folgenden theoretischen Ausführungen sollen einen Überblick über den Gegenstand der Evaluation sowie den aktuellen Stand der Evaluationsforschung geben. Zudem soll in einem zweiten Teil spezifisch auf die Eigenheiten von Evaluation im Bereich E-Learning eingegangen werden, da diese im Gesamtkontext dieser Arbeit im Fokus steht.

2.1.2 Begriffsbestimmung

Evaluation leitet sich vom lateinischen Begriff valor (Wert) und der Vorsilbe e / ex (aus) ab und bedeutet übersetzt „Bewertung“ oder „einen Wert aus etwas ziehen“ (Meyer & Höhns, 2002). Evaluation wird alltagssprachlich auch mit Bewertung oder Beurteilung gleichgesetzt. Eine allgemeingültige, wissenschaftliche Definition von Evaluation existiert jedoch nicht. Verwandte Begriffe und Konzepte, welche teilweise synonym, teilweise im Sinne einer speziellen Form von Evaluation verwendet werden, sind zum Beispiel Begleitforschung, Qualitätskontrolle oder Wirkungskontrolle (Tergan, 2000). Evaluation umfasst eine Vielzahl an Verhaltensweisen und entzieht sich damit nach Ansicht einiger Autoren einer abstrakten, die Wirklichkeit umfassenden Definition. So ist es nach Guba und Lincoln (1989, S. 21) nicht möglich, Evaluation begrifflich eindeutig festzulegen: „We take definitions of evaluation to be human mental constructions ... There is no answer to the question, ‘But what is evaluation really?’ and there is no point in asking it.”

Vielmehr hat sich in der Zwischenzeit zur Ausdifferenzierung des Begriffs wie auch aus didaktischen Gründen eine gewisse Vielfalt von Definitionen etabliert (Wottawa & Thierau, 1998). Nach Böttcher (2006) wird unter Evaluation die Methode systematischer Datensammlung, die Analyse und eine an Kriterien orientierte Bewertung der Befunde mit dem primären Ziel, Impulse für die Verbesserung von Massnahmen und Systemen zu liefern, verstanden. Evaluation ist dabei direkt an Qualitätssicherung und –entwicklung gekoppelt. Evaluation steht also nicht nur für ein spezifisches Handeln, das sowohl die Gewinnung von Informationen als auch deren Bewertung zum Ziel hat, sondern auch für das Ergebnis dieses Prozesses. Im Unterschied zu fachbezogener, wissenschaftlicher Forschung stellt Evaluation keinen Selbstzweck dar, sondern soll immer einen Nutzen stiften (Stockmann, 2006).

Gemäss Fricke (2002) zeichnet sich Evaluation vor allem durch eine über die reine Erhebung und Quantifizierung von Variablen hinausgehende, umfassende Bewertung aus. Ähnlich versteht Rindermann (2001) Evaluation als eine systematische Analyse und empirische Untersuchung von Konzepten, Bedingungen, Prozessen und Wirkungen zielgerichteter Aktivitäten zum Zweck ihrer Bewertung und Modifikation. Eine blosser Beschreibung und Messung transzendierend, besteht das Ziel von Evaluation in Bewertung und Optimierung. Weitere Definitionsversuche sind in der Tab. 1 aufgeführt.

Tab. 1: Verschiedene Definitionsansätze von Evaluation.

| Definitionsansatz | Autor |
|--|--|
| Rindermann definiert Evaluation als „eine Form rationaler Problemlösungsstrategie, die für Entscheidungen zuverlässige empirische Datengrundlagen liefert sowie Entscheidungen mittels Klärung von Zielen und Dekomposition von komplexen Problemen rationaler gestalten möchte“. Gegenstand der Evaluation sind Programme, Verfahren, Produkte, Institutionen und Tätigkeiten von Personen. | Rindermann (2003, S. 233) |
| Gemäss Beywl zielt Evaluation auf die Verbesserung gesellschaftlicher Praxis, indem sie Programme, Massnahmenbündel oder Materialien systematisch, d.h. methodisch angeleitet und an Gütekriterien überprüfbar, beschreibt und bewertet. | Beywl (1988) |
| Das Joint Committee definiert Evaluation als die systematische Untersuchung der Verwendbarkeit oder Güte eines Gegenstandes. | Joint Committee on Standards for Educational Evaluation (1999) |

| | |
|--|----------------------------|
| „Evaluation is the process of determining the merit, worth and value of things, and evaluations are the products of the process.“ | Scriven (1991, S. 139) |
| Vedung beschränkt den Begriff Evaluation auf die Beurteilung öffentlicher Interventionen. Evaluation ist demnach die „retrospektive Bewertung des Verdienstes, der Güte und des Wertes der Implementation, der Leistung und der Ergebnisse staatlicher Interventionen mit dem Zweck, in zukünftigen praktischen Handlungen eine Rolle zu spielen.“ | Vedung (1999, S. 2) |
| Gemäss Baumgartner meint Evaluation alle Aktivitäten und/oder Ereignisse, welche die Bedeutung, Verwendbarkeit, (Geld-)Wert, Wichtigkeit oder Zweckmässigkeit einer Sache beurteilen, bzw. bewerten. | Baumgartner (1997) |
| In sozialen Kontexten synonym zu Evaluation verwendete Begriffe: Erfolgskontrolle, Effizienzforschung, Begleitforschung, Bewertungsforschung, Wirkungskontrolle, Qualitätskontrolle. | Wottawa und Thierau (1998) |

Um das Problem der bestehenden Definitionsvielfalt zu umgehen, beschränken sich Wottawa und Thierau (1998) auf die Festlegung allgemeiner Kennzeichen von Evaluation:

- a) Sämtliche Tätigkeiten im Zusammenhang mit Evaluation haben etwas mit „bewerten“ zu tun.
- b) Evaluation ist ziel- und zweckorientiert und beabsichtigt, praktische Massnahmen zu überprüfen, zu verbessern oder über sie zu entscheiden.
- c) Evaluationsmassnahmen sollten dem aktuellen Stand wissenschaftlicher Technik und Forschungsmethoden angepasst sein.

Mittels dieser Kriterien versuchen Wottawa und Thierau die wissenschaftliche Evaluation von anderen Vorgehensweisen abzugrenzen. Grundlegend lässt sich festhalten, dass - unabhängig vom methodischen Ansatz - Evaluation systematische Bewertungen beinhaltet, welche auf systematischen Untersuchungen basieren.

2.1.3 Sichtweisen und Anwendungsbereiche von Evaluation

Die historische Entwicklung der Evaluationsforschung hat unterschiedliche Standpunkte hervorgebracht. Baumgartner (1997) unterscheidet dabei folgende Sichtweisen:

Quantitative Sichtweise:

Evaluation beschränkt sich auf die Konstruktion und Auswertung von Tests und quantitativen Verfahren.

Methodische Sichtweise:

Evaluation bedeutet Methodenlehre und beschränkt sich damit auf die Beschreibung, Diskussion und Umsetzung quantitativer wie auch qualitativer Methoden.

Pragmatische Sichtweise:

Evaluation zielt auf praktische Massnahmen zur Optimierung ab und schliesst damit beispielsweise vergleichende Untersuchungen aus.

Sichtweise der Sozialforschung:

Evaluation wird einfach als eines von vielen Gebieten der Sozialforschung verstanden.

Bewertungssichtweise:

Was Evaluationen letztlich ausmacht, ist der Bewertungsprozess (Festlegen des Wertes einer Sache) selber.

Relativistische Sichtweise:

Auf einen Definitionsversuch darüber, was Evaluation ist, wird verzichtet. Entscheidend sind bei dieser Position die Unterschiede in Ansatz und Praxis.

Jede dieser Sichtweisen führt zu einem spezifischen Evaluationsverständnis, woraus unterschiedliche Schwerpunkte, Herangehensweisen und Interpretationsansätze resultieren. Evaluation kann weiter als Prozess oder als Produkt verstanden werden (Baumgartner, 1997): Als Prozess ist Evaluation Teil der angewandten Forschung (Evaluationsforschung), welche zum einen handlungsanleitenden Charakter in Form einer Entscheidungshilfe besitzt und zum anderen auf Erkenntnisgewinn ausgerichtet ist. Trotz des ausgeprägten Praxisbezugs und pragmatischen Ansatzes beinhaltet sie aber auch Grundlagenforschung, Theoriebildung und Metaevaluation. Als Produkt wird das Resultat

tat der Forschung in einem bestimmten Evaluationsfeld wie auch die Disziplin selber verstanden.

Typische Gegenstände, mit denen sich Evaluationen beschäftigen, sind gemeinhin Bildungs- oder soziale Interventionsprogramme. In einem weiteren Sinn beschäftigt sich Evaluationsforschung jedoch mit einer Vielzahl anderer Objekte (Bortz & Döring, 2002; Mertens & McLaughlin, 2004):

- Personen (z.B. Evaluation von Hochschuldozenten durch Studierende)
- Umweltfaktoren (z.B. Auswirkungen von Fluglärm)
- Produkte (z.B. vergleichende Analysen der Wirkung von Psychopharmaka)
- Techniken und Methoden (z.B. Trainingsmethoden für Hochleistungssportler)
- Zielvorgaben (z.B. Vergleich der Ausbildungsziele „Fachkompetenz“ oder „soziale Kompetenz“ bei einer Weiterbildungsmaßnahme)
- Projekte (z.B. Evaluation eines universitären E-Learning-Angebots)
- Systeme und Strukturen (z.B. Evaluation der Auswirkung verschiedener Unternehmensstrukturen auf die Mitarbeiterzufriedenheit)
- Forschung (z.B. Gutachten über Forschungsanträge)

Der breite Anwendungsbereich von Evaluationsforschung zeigt, dass sämtliche forschende Tätigkeitsfelder tangiert sind, welche die Überprüfung und Bewertung von Massnahmen oder Interventionen zum Inhalt haben.

2.1.4 Abgrenzung zu verwandten Konzepten

Ziel und Orientierungsrahmen für die Anwendung von Ansätzen und Methoden der Evaluationsforschung bildet die Qualitätssicherung (Reinmann-Rothmeier, Mandl & Prenzel, 1997). Qualitätssicherung wird hier als ein der Evaluation übergeordnetes Konzept aufgefasst und die Evaluation als ein Instrument zur Sicherung und Kontrolle von Qualität verstanden.

Ein umfassendes System der Qualitätssicherung repräsentiert das Konzept des Total Quality Managements (TQM). Der TQM-Ansatz beinhaltet Qualitätssicherung, -management als auch -entwicklung und zielt darauf ab, Qualität als Systemziel festzusetzen und dadurch dauerhaft zu garantieren. Gemeinsam ist TQM und Evaluationskonzepten das Streben nach Qualitätsverbesserung. Im Unterschied zur Evaluation kommt TQM jedoch vor allem in profitorientierten Unternehmungen zur Anwendung, während Evaluation eher in öffentlichen Einrichtungen und Non-Profit-Organisationen genutzt wird. Der Kontext und die Struktur, in dem die beiden Konzepte verwendet werden, unterscheidet sich deutlich, was dazu führt, dass diese auch andere wissenschaftstheoretische Wurzeln aufweisen.

TQM-Konzepte haben ihren Ursprung im betriebswirtschaftlichen Bereich. Evaluation ist dagegen stark sozialwissenschaftlich ausgerichtet, was im Vergleich zu TQM zu unterschiedlichen Perspektiven und Ansätzen führt (Stockmann, 2006). Controlling ist ein weiteres Konzept, welches mit Evaluation und Qualitätssicherung verwandt ist. Das Controlling ist aber in erster Linie auf Fragen von Kosten und Nutzen ausgerichtet. Die folgende Übersicht soll das Gesagte in Anlehnung an Behrmann (2006) zusammenfassen und verdeutlichen:

Evaluation:

Die Evaluation ist empirisch-sozialwissenschaftlich orientiert. Sie untersucht in erster Linie personenbezogen soziale Aspekte sowie Effekte von Bildung. Bei der Evaluation geht es somit meist im Sinne eines Monitoring um die systematische und wissenschaftsorientierte Begleitung von Aktivitäten im sozialen Kontext.

Qualitätssicherung, Qualitätsmanagement und Qualitätsentwicklung:

Hierbei handelt es sich um Managementinstrumente, welche ihren Ursprung im betriebswirtschaftlichen Bereich haben. Sie dienen vor allem der Kontrolle und der Optimierung von Leistung bei gegebenen Zielen. Als TQM bezieht sich dieser Ansatz in einem umfassenden Sinn auf die gesamte Organisation oder Unternehmung und beinhaltet neben der operativen Umsetzung auch die strategische Orientierung.

Controlling:

Controlling ist primär auf das Verhältnis von Kosten und Nutzen ausgerichtet. Es dient in diesem Sinne der Überprüfung der Wirtschaftlichkeit anhand von Kennzahlen. Controlling kann in einem umfassenderen Sinn auch strategisch ausgerichtet sein und der Steuerung der gesamten Organisation basierend auf ihren Zielen dienen.

Evaluation ist zwar sozialwissenschaftlich orientiert, unterscheidet sich von sozialwissenschaftlicher Forschung aber dadurch, dass bei der Evaluation die Zuweisung von Werten zentral ist, während die Sozialforschung der Wertfreiheit verpflichtet ist. In den Sozialwissenschaften sind die Problemdefinition sowie die Verwertung der Ergebnisse aus dem Forschungszusammenhang ausgeklammert. Im Fall der Evaluation ist dies gerade umgekehrt: die Interessen und Ziele von Auftraggebern und Beteiligten sind hier entscheidend (Baumgartner, 2005). Tab. 2 verdeutlicht die Unterschiede zwischen Sozialforschung und Evaluation.

Tab. 2: Unterschiede zwischen Evaluation und Sozialforschung (Baumgartner 2005).

| Evaluation | | Sozialforschung |
|---------------------------|-----|---|
| Bewertend | < > | Wertfrei |
| Entscheidungsorientiert | < > | Schlussfolgerungsorientiert |
| Verbesserungsorientiert | < > | Erklärungsorientiert |
| Wert | < > | Wahrheit |
| Gesellschaftlicher Nutzen | < > | Empirische Überprüfbarkeit; logische Konsistenz |
| Evaluator | < > | Wissenschaftler |
| Auftrag | < > | Forschung |
| Interessenorientiert | < > | Erkenntnisorientiert |
| Evaluationsumfeld | < > | Scientific Community |
| Oft nicht generalisierbar | < > | Generalisierbar; theoriebildend |

Evaluation bedeutet problemorientiertes Handeln und kann damit grundsätzlich nicht auf den Erkenntnissen einer einzelnen Wissenschaft aufbauen. Die Fragestellungen, mit denen sich die Evaluationsforschung konfrontiert sieht, erfordern ein fächerübergreifendes Wissen, welches um Praxiswissen ergänzt wird (Wottawa & Thierau, 1998). Oft bleibt Forschung aber an eigenen Fächertraditionen statt an spezifischen Fragestellungen orientiert. Forschungsansätze, Theorien und Methoden anderer Disziplinen werden übersehen (Rindermann, 2003). Nur ein interdisziplinärer Ansatz führt aber zu Lösungen, welche den Problemen und Anforderungen in der Praxis gerecht werden. Für die Evaluationsforschung ist es daher elementar, dass ein Erfahrungsaustausch von Evaluations-Know-how zwischen den einzelnen Anwendungsfeldern stattfindet.

2.1.5 Historische Dimension

Die systematische Evaluationsforschung ist ein Produkt des 20. Jahrhunderts. Die Anwendung von sozialwissenschaftlicher Methodik im Rahmen von Evaluation fällt zusammen mit dem vermehrten Einsatz und der Ausdifferenzierung von Forschungsmethoden im Allgemeinen sowie mit ideologischen, politischen als auch demografischen Veränderungen während dieser Zeit (Rossi, Lipsey & Freeman, 2004). Im Zuge der Zunahme des Wissens in den empirischen Wissenschaften entwickelte sich eine spezifische Evaluationsforschung. Ausgehend von den USA bekam diese einen immer höheren Stellenwert und wurde in nahezu allen gesellschaftlichen Bereichen (Bildung, Wirtschaft, Politik, Verwaltung, Umwelt, Gesundheit, Militär) zu einer wesentlichen Gestaltungshilfe. Die Gesellschaften realisierten, dass sich ihr Handeln ebenfalls unter Optimierungsas-

pekten rational und zielorientiert gestalten liess (Wottawa & Thierau, 1998). Beywl (1988) beschreibt für die Entwicklung der Evaluation in den USA bis in die 1990er Jahre sechs Stufen (vgl. auch Madaus & Stufflebeam, 2000):

1. Akkreditation versus Testverfahren (bis 1925)

Zu Beginn des 20. Jahrhunderts entstanden zwei Ansätze, welche den bis heute dauernden methodologischen Grundkonflikt repräsentieren. Zum einen das Akkreditationsmodell, wonach Schulen von Experten anhand von qualitativen Kriterien beurteilt wurden. Zum anderen das Testmodell, bei denen Schülerleistungen anhand von objektiv messbaren Grössen und mittels standardisierten Messinstrumenten erfasst wurden. Im Akkreditationsansatz und im Testmodell zeigen sich zwei konkurrierende Modelle des Bewertungsansatzes: die Messung mittels expliziter, quantifizierbarer Kriterien, der die Beurteilung anhand subjektiver, impliziter und qualitativer Kriterien gegenübersteht.

2. Curriculumevaluation (1925-1945)

Diese Phase war stark durch Ralph W. Tyler geprägt. Er zeichnete sich an der Ohio State University zuständig für die Verbesserung von Lehre und Lernen. Dabei entwickelte er einen stark an wissenschaftlichen Kriterien angelehnten Ansatz, bei dem die Leistungsmessung durch die Messung des Lernerfolgszuwachses ersetzt wurde. Das Ziel bestand in einer fortlaufenden und begleitenden Verbesserung von pädagogischen Prozessen. Tyler gilt allgemein als geistiger Vater der Programmevaluation.

3. Stagnation (1945-1957)

Die Nachkriegszeit war durch einen Verlust an Interesse an Evaluationsmethodologie gekennzeichnet. Zum einen wurde Tylers Ansatz als zu aufwändig und komplex erlebt. Zum anderen gab es im Zuge des Wirtschaftswachstums nur wenig Interesse an Verantwortlichkeit für pädagogische Qualitäten. Die methodologische Diskussion beschränkte sich auf Fragen der Testkonstruktion, der Auswertung grosser Datenmengen etc.

4. Expansion und Innovation (1957-1970)

Infolge des durch das russische Raumfahrtprogramm (Sputnik) ausgelösten Schocks – die USA sah sich wissenschaftlich-technisch ins Hintertreffen geraten – wurden die finanziellen Mittel für Evaluationsmassnahmen massiv aufgestockt. Folge davon war unter anderem, dass an den Universitäten Evaluationslehrstühle eingerichtet sowie erste Publikationsorgane geschaffen wurden. Mitte der 60er-Jahre wurden dank des Reformprogramms zur Verbesserung der Chancengleichheit in der „Great Society“

zusätzliche Mittel gesprochen. In dieser Zeit fand auch eine angeregte methodologische Diskussion statt, welche der Praxis weit voraus war.

5. Verunsicherung (1970-1980)

In den 1970er-Jahren wurde ernüchtert festgestellt, dass die meisten Evaluationsergebnisse von den Entscheidungsträgern in Politik und Verwaltung nicht genutzt wurden. Dies löste wiederum eine intensive methodologische Debatte aus. Gefordert wurden eine verstärkte Zusammenarbeit zwischen Evaluatoren und Entscheidern sowie noch mehr methodische Strenge in der Planung und Durchführung von Untersuchungen.

6. Herausforderung (1980-1988)

Unter Präsident Reagan erlebte die Sozial- und Bildungspolitik in den USA einen starken Umschwung. Sozial- und Bildungsprogramme wurden grösstenteils für überflüssig erklärt und benötigten damit auch keine Beschreibung und Bewertung mehr. Der Etat für Evaluationsmassnahmen erlebte in der Folge eine starke Kürzung. Der Ruf nach stärkerer Professionalisierung, nach effektiven und effizienten Verfahren, nach einem marktgerechten Auftreten der Sozial- und Erziehungswissenschaften wurde lauter.

Die Forderung nach mehr Professionalität und stärkerer Institutionalisierung führte 1986 zur Gründung der American Evaluation Association (AEA)⁹. Das Ziel der Organisation besteht in der Verbesserung der Evaluationspraktiken und –methoden, in der Förderung von Evaluation als beruflicher Tätigkeit und des Vorantreibens der Evaluationsforschung. In Europa setzte die Professionalisierung der Evaluationsforschung angeregt durch die Entwicklungen in den USA Ende der 1960er-Jahre ein. Der Schwerpunkt der Evaluationen lag hier wie in den USA in erster Linie im Nachweis der Effektivität von Programmen in der Bildungs-, Gesundheits- und Sozialpolitik (Meyer & Höhns, 2002). Auch hier wurden in verschiedensten Ländern nach amerikanischem Vorbild Evaluationsgesellschaften gegründet.¹⁰ Die Entwicklungsforschung hat sich mittlerweile auch in Westeuropa etabliert und expandiert weiter. „Ohne systematische wissenschaftliche Beobachtungen und Beurteilungen der Wirkung politischen Handelns kommen moderne Demokratien nicht mehr aus“ (Meyer & Höhns, 2002, S. 6). Ein Beispiel hierfür ist das Programme for International Student Assessment (PISA), welches den Vergleich von

⁹ American Evaluation Association: <http://www.eval.org>

¹⁰ European Evaluation Society: <http://www.europeanevaluation.org>
Deutsche Gesellschaft für Evaluation: <http://www.degeval.de>
Schweizerische Evaluationsgesellschaft: <http://www.seval.ch>

Schulleistung verschiedener europäischer Länder zum Ziel hat und international grosse Beachtung findet.

2.1.6 Evaluationsrichtlinien

Obwohl keine eigentliche Berufsgattung existiert, welche sich der Evaluation verschrieben hat, so hat sich die Evaluationsforschung seit den 70er Jahren des 20. Jahrhunderts doch stark professionalisiert. Dementsprechend wurde das Verlangen nach Richtlinien für die Evaluationstätigkeit laut (Rossi et al., 2004). In den USA wurden in den 1990er Jahren diesbezüglich zwei Vorstösse unternommen: die American Evaluation Association erliess 1995 die *Guiding Principles for Evaluators*. Diese enthält fünf generelle Prinzipien, welche bei der Evaluation berücksichtigt werden sollten:

- a) Evaluation basiert auf einer systematischen, datenbasierten Vorgehensweise.
- b) Evaluatoren bieten ausreichend Fachkompetenz zur Durchführung von Studien.
- c) Integrität und Ehrlichkeit sind grundlegend für den gesamten Evaluationsprozess.
- d) Evaluatoren behandeln die Personen, mit denen sie im Zug ihrer Tätigkeit in Kontakt kommen, mit angemessenem Respekt.
- e) Evaluatoren berücksichtigen die unterschiedlichen Interessen und Werte, welche im Zusammenhang mit dem Gemeinwohl stehen.

Im Gegensatz zu diesen eher allgemein gehaltenen Prinzipien publizierte das Joint Committee on Standards for Educational Evaluation 1994 die *Program Evaluation Standards*, welche in detaillierter Form eine grosse Bandbreite an Aspekten umfasst, welche es während des Evaluationsprozesses zu berücksichtigen gilt. Die Standards sind den vier zentralen Attributen einer Evaluation zugeordnet (in Anlehnung an Sanders, 1999):

Utility (Nützlichkeit)

Nützlichkeitsstandards legen fest, ob eine Evaluation den tatsächlichen Evaluationsbedürfnissen der jeweiligen Adressaten gerecht wird.

Feasibility (Durchführbarkeit)

Durchführbarkeitsstandards verlangen, dass die Evaluationen realistisch, gut durchdacht, diplomatisch und wirtschaftlich durchgeführt werden sollen.

Propriety (Korrektheit)

Korrektheitsstandards fordern, dass Evaluationen in einem ethisch und rechtlich korrekten Rahmen stattfinden.

Accuracy (Genauigkeit)

Die Genauigkeitsstandards sollen sicherstellen, dass Evaluationen präzise Informationen über die Verwendbarkeit und Güte eines Programms offenlegen und vermitteln.

Die Entwicklung und Verbreitung solcher Standards und Richtlinien hat massgeblich zu einer gesteigerten Qualitätssicherung von Evaluationen geführt. Die Standards dienen auch in Europa als Vorbild zur Gestaltung eigener Richtlinien.

2.1.7 Spezifische Aspekte von Evaluation

Abhängig von der Phase im Projektverlauf kommen unterschiedliche Evaluationsassumtionen mit unterschiedlicher Zielsetzung zum Einsatz. Je nachdem wer die Massnahmen durchführt, spricht man von interner oder externer Evaluation. Gegenstand der Evaluation kann zudem entweder ein bestimmtes Produkt oder der Prozess der Produktentwicklung sein.

2.1.7.1 Arten von Evaluation

Evaluationsmassnahmen bestimmen sich nach dem Zweck, dem die Massnahmen dienen und dem Zeitpunkt im Projektverlauf, an dem sie erfolgen. Grundsätzlich lassen sich zwei Arten von Evaluation unterscheiden: formative und summative Evaluation (Wottawa & Thierau 1990)¹¹.

Formative Evaluation

Der primäre Zweck der formativen Evaluation besteht in der Qualitätssicherung. Sie erfolgt entwicklungsbegleitend und dient in erster Linie der Ermittlung von Schwachstellen. Die formative Evaluation ist aktiv-gestaltend, prozessorientiert und konstruktiv angelegt und liefert damit Daten zur Optimierung. Bei den hierbei eingesetzten Methoden handelt es sich meist um solche mit informellem Ansatz, die schnell relevante Daten lie-

¹¹ Teilweise wird von einer dritten Art von Evaluation, der ex-ante Evaluation, gesprochen (Stockmann, 2000). Dabei handelt es sich um Machbarkeitsstudien, welche der Untersuchung der Ausgangs- und Rahmenbedingungen dienen. Das Ziel solcher Evaluationen besteht in einer Einschätzung möglicher Durchführungsprobleme sowie der Abschätzung der Erfolgchancen von geplanten Massnahmen.

fern (Tergan, 2000). Formative Evaluationen werden vor allem bei der Entwicklung und Implementierung neuer Massnahmen eingesetzt und sind meistens erkundend angelegt.

Summative Evaluation

Die summative Evaluation dient in erster Linie der Kontrolle von Qualität, Wirkung und Nutzen eines Produkts. Diese Form der Evaluation findet nach Abschluss der Implementationsphase statt und zeigt, ob ein Produkt oder Programm die beabsichtigte Wirkung auch erzielt und den gestellten Ansprüchen genügt. Normalerweise werden in der Evaluationsforschung Hypothesen geprüft, nachdem eine Massnahme abgeschlossen ist.

Evaluationsmethoden lassen sich nicht immer strikt einer der beiden Evaluationsarten zuordnen. So kann beispielsweise eine Akzeptanzbefragung während einer Pilotphase zur Qualitätssicherung (formative Evaluation), aber auch nach der Implementation zur Qualitätskontrolle (summative Evaluation) eingesetzt werden.

2.1.7.2 Produkt- und Prozessevaluation

Evaluationsansätze lassen sich nach der Art des Inhalts der Analyse einordnen. Inhalt kann ein Produkt, aber auch ein Prozess sein, so dass man Produktevaluation von Prozessevaluation unterscheidet.

Produktevaluation

Wenn sich Evaluation „auf ein entwickeltes Produkt, wie ein Bildungsangebot oder Teile davon bezieht und Aspekte der Qualität, der Wirkungen (z.B. Lernerfolg), der Effizienz und des Nutzens im Vordergrund stehen“ spricht man von Produktevaluation (Tergan, 2000, S. 27). Diese erfolgt meist nach abgeschlossener Entwicklungsphase, kann aber auch bereits während des Entwicklungsprozesses erfolgen. Im Fall der Evaluation eines Lehrgangs wird beispielsweise häufig die Wirkung anhand des Lernerfolgs erfasst.

Prozessevaluation

Die Prozessevaluation befasst sich mit Aspekten des Planungs- und Entwicklungsprozesses. Der Durchführungsprozess selber wird dabei erfasst und analysiert. Das Ziel liegt dabei unter anderem darin, Schwachstellen im Verfahren der Planung und Abstimmung zwischen den einzelnen an der Durchführung der Evaluation beteiligten Personen aufzudecken. Im Sinne einer Metaevaluation kann sich Prozessevaluation auch auf die Durchführung der Evaluation selber beziehen (Tergan, 2000).

2.1.7.3 Interne versus externe Evaluation

Ein weiterer Aspekt, in dem sich Evaluationen unterscheiden können, ist die Stellung der evaluierenden Person im Entwicklungsprozess.

Interne Evaluation

Interne Evaluationen kennzeichnen sich dadurch, dass sie von derselben Organisation durchgeführt werden, welche auch mit der Entwicklung des zu evaluierenden Produkts oder der zu evaluierenden Massnahme betraut ist. Wird die interne Evaluation gleichzeitig von derselben funktionalen Organisationseinheit durchgeführt, welche auch die operative Tätigkeit inne hat, spricht man von Selbstevaluation (Stockmann, 2006). In diesem Fall ist die evaluierende Person direkt auch an der Entwicklung beteiligt. Der Vorteil einer internen Evaluation liegt darin, dass die Evaluatoren über grosse Sachkenntnis verfügen und gezielter vorgehen können. Der Nachteil liegt in der Gefahr des Mangels an kritischer Distanz und Unabhängigkeit sowie in häufig fehlendem Methodenwissen.

Externe Evaluation

Bei einer externen Evaluation wird die Evaluation von Personen durchgeführt, die von Geldgebern und der mit der Entwicklung der zu evaluierenden Massnahme vertrauten Organisation unabhängig sind. Der Vorteil der externen Evaluation liegt in der Unabhängigkeit der Evaluatoren, deren profunden Methodenkenntnissen und ihrem professionellen Evaluationswissen (Stockmann, 2006). Nachteile können sich in einem Mangel an Bewusstsein der tatsächlichen Problemlage sowie in einer Abwehrreaktion der evaluierten Personen zeigen.

Oftmals werden interne und externe Evaluation kombiniert eingesetzt, um dadurch von den Vorteilen beider Vorgehensweisen profitieren zu können.

2.1.8 Evaluation im Projektverlauf

Evaluation kann grundsätzlich in sämtlichen Projektphasen Anwendung finden. In der Planungsphase besteht das Ziel von Evaluation darin, Daten zu erheben, die für die Verbesserung der Projektkonzeption genutzt werden können. In der Implementationsphase steht die Qualitätssicherung unter Berücksichtigung der gewählten Ziele, Qualitätskriterien und Rahmenbedingungen im Fokus, während in der Wirkungsphase die Evaluation der Wirkung (Akzeptanz, Kosten-Nutzen-Verhältnis, Vergleich mit alternativen Massnahmen) einer Massnahme oder eines Produktes geprüft werden (Tergan, 2000). Je nach Projektphase hat die formative oder die summative Perspektive grössere Bedeutung. So gibt es in der Planungs- und Designphase kaum Ansatzpunkte für summative Evaluation. Während der Durchführungsphase sind formative als auch summative Evaluation mög-

lich. Auf Grund des fehlenden Gestaltungsaspekts in der Wirkungsphase kommt hier in der Regel summative Evaluation zum Einsatz. Da deren Ergebnisse jedoch für Folgeprojekte Bedeutung haben können, bekommen sie auch formativen Charakter (Stockmann, 2006).

Tab. 3: Evaluationskonzepte im Projektverlauf (Stockmann, 2006).

| Projektphasen | Analyseperspektive | Evaluationskonzept |
|--------------------------------------|--------------------|---|
| Programmformulierung / Planungsphase | ex-ante | Formativ: aktiv gestaltend, prozessorientiert, konstruktiv |
| Implementationsphase | on-going | Formativ sowie summativ |
| Wirkungsphase | ex-post | Summativ: zusammenfassend, bilanzierend, ergebnisorientiert |

2.1.9 Funktionen von Evaluation

Die Funktionen der Evaluation können vielfältiger Natur sein. Rowntree (1992) unterscheidet drei Hauptfunktionen: eine strategisch-politische Funktion, eine Kontroll- und Entscheidungsfunktion sowie eine Erkenntnisfunktion¹².

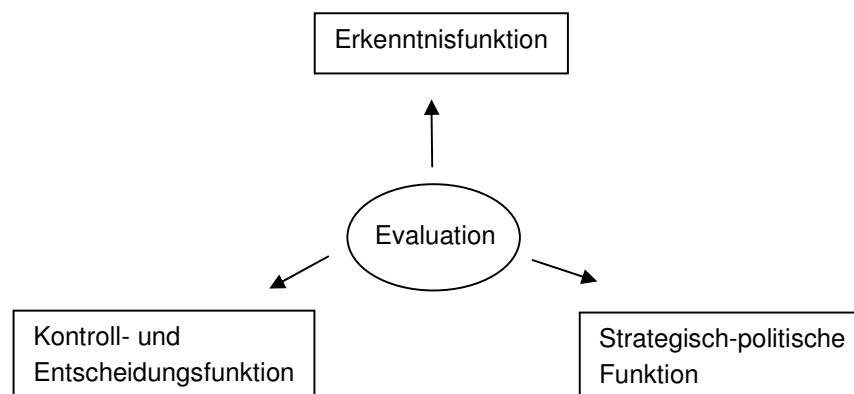


Abb. 11: Funktionen von Evaluation (in Anlehnung an Rowntree, 1992).

Im Sinne einer Kontroll- und Entscheidungsfunktion wird Evaluation als längerfristiger Prozess verstanden, bei dem es darum geht, laufend zu kontrollieren und Entscheidungen

¹² Stockmann (2006) erwähnt Entwicklung als eine weitere Funktion von Evaluation. Dabei sieht er Evaluation als einen Faktor, der Transparenz und Dialogmöglichkeiten schafft, was wiederum die Entwicklung vorantreibt.

zu treffen mit dem Ziel zu optimieren. Damit wird die Qualitätssicherung im Entwicklungsablauf implementiert, was zu einem steten Optimierungsprozess führt.

Als strategisch-politische Funktion soll Evaluation die Wirkung respektive den Nutzen einer Massnahme aufzeigen und diese letztlich legitimieren. Die Evaluationsergebnisse richten sich dabei in erster Linie an aussenstehende Personen oder Institutionen. So wird beispielsweise den Nutzern eines Bildungsprogramms dessen Wert aufgezeigt oder den finanziell fördernden Institutionen dargelegt, wie sinnvoll die erhaltenen Gelder eingesetzt werden.

Evaluation als Erkenntnisfunktion wiederum bedeutet, dass Evaluationen Daten liefern, auf deren Basis Erkenntnisse über Effekte von Massnahmen resultieren. Damit lassen sich Planungs- und Steuerungsentscheide auf eine rationale Grundlage stellen.

Die einzelnen Funktionen sind nicht unabhängig voneinander zu verstehen. Welche der Funktionen von primärem Interesse ist und am stärksten gewichtet wird, hat jedoch Auswirkung auf die grundsätzliche Herangehensweise und auf die Konzeption sowie die Durchführung einer Evaluation (Stockmann, 2006). Bei einem pragmatischen Ansatz wird die Kontroll- und Entscheidungsfunktion im Vordergrund stehen, während bei Evaluation als wissenschaftlicher Begleitforschung die strategisch-politische Funktion sowie die Erkenntnisfunktion höher gewichtet sein dürften.

2.1.10 Evaluationsprozess und Evaluationsmethoden

Der Evaluationsprozess gestaltet sich im Normalfall als ein planvoller, regelgeleiteter Prozess. Anhand einer zweckgerichteten Auswahl an Kriterien wird eine zu evaluierende Massnahme systematisch geprüft und bewertet. Schliesslich soll die Bewertung eine Kommunikation ermöglichen mit dem Ziel, Konsequenzen abzuleiten (Abs, Maag-Merki & Klieme, 2006). Da die Ergebnisse einer Evaluation wiederum die Voraussetzungen schaffen, in denen neue Evaluationen sinnvoll sein können, kann dieser Prozess auch als zyklisch gedacht werden (Abb.12).

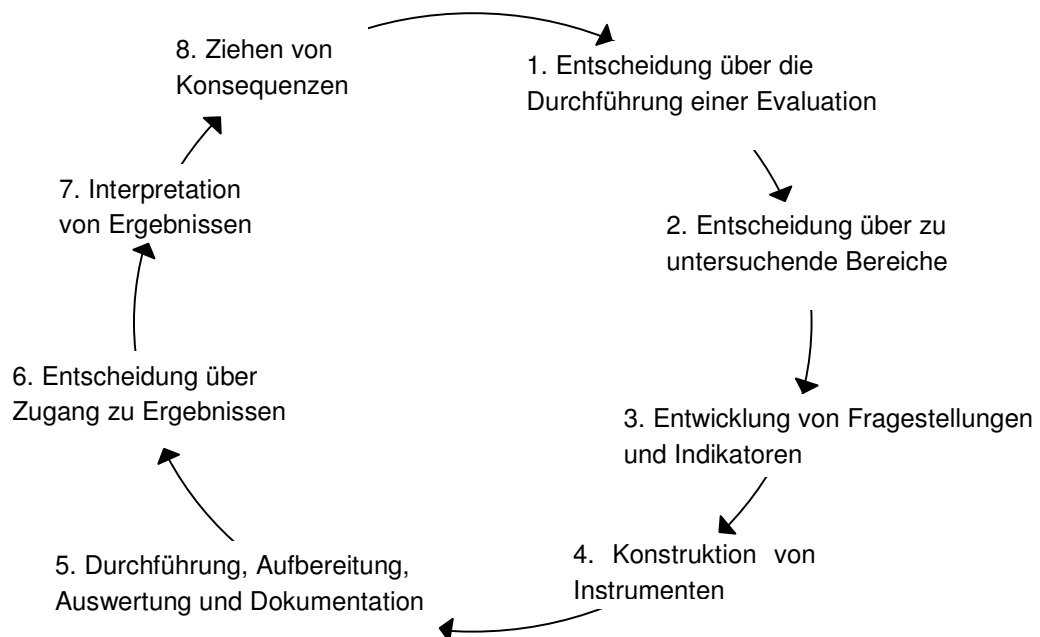


Abb. 12: Evaluationsprozess (in Anlehnung an Abs, Maag Merki & Klieme, 2006)

Ein alternatives Vorgehen, welches den Schwerpunkt auf den Bewertungsaspekt legt, stammt von Scriven (1991). Dabei wird Evaluation als vierstufiger Prozess verstanden:

1. Festlegung von Wertekriterien: In einer ersten Phase werden Kriterien bestimmt, welche vom Evaluanden erfüllt werden müssen.
2. Bestimmung von Leistungsstandards: Für die einzelnen Kriterien werden im Zuge einer Operationalisierung Leistungsstandards bestimmt, welche es vom Evaluanden zu erreichen gilt.
3. Messung und Vergleich: Die definierten Kriterien werden beim Evaluanden gemessen und anhand der Leistungsstandards auf Zielerreichung geprüft.
4. Werturteil: In der letzten Phase des Evaluationsprozesses müssen die Ergebnisse interpretiert werden und zu einem einheitlichen Werturteil integriert werden.

Evaluationsmethoden sind die Werkzeuge, mit denen sich die Daten und Informationen generieren lassen, die der eigentlichen Bewertung zugrunde liegen. Evaluation im wissenschaftlichen Kontext unterscheidet sich von Evaluation im Alltagshandeln grundsätzlich dadurch, dass empirische Methoden zur Informationsgewinnung und systematische Verfahren zur Informationsbewertung anhand transparent gemachter Kriterien verwendet werden, was eine intersubjektive Nachprüfbarkeit ermöglicht (Stockmann, 2006). Die Methoden lassen sich allgemein in quantitative und qualitative Ansätze unterteilen und decken die gesamte Bandbreite wissenschaftlichen Arbeitens ab. Je nach Gegenstand und Evaluationsziel können die zu wählenden Methoden sehr unterschiedlich ausfallen. Methoden, die im Rahmen von Evaluationen häufig zur Anwendung kommen, sind: Befragung, Beobachtung, Inhaltsanalysen sowie empirische Untersuchungen (Bortz & Döring, 2002).

Oft besteht die Gefahr, dass sich die Wahl der Methode in erster Linie nach den zeitlichen, monetären oder personellen Gegebenheiten ausrichtet. Die Wahl des methodischen Ansatzes sollte sich jedoch möglichst an Sachaspekten orientieren, auch wenn der Handlungsspielraum auf Grund beschränkter Ressourcen oftmals eingeschränkt ist (Wottawa & Thierau, 1998).

2.1.11 Evaluationsformen und Interventionsebenen

Stufflebeam (2000) definiert in seinem CIPP-Modell vier verschiedene Evaluationsformen. Jede Evaluationsform entspricht dabei einer bestimmten zeitlichen Phase im Evaluationsprozess:

Kontextevaluation (Context):

Die Kontextevaluation behandelt die Frage nach den Rahmenbedingungen der Evaluation. Hierbei geht es somit um die Bestimmung des Bedarfs, der Probleme sowie der Ressourcen in einem gegebenen Kontext.

Inputevaluation (Input):

Die Inputevaluation dient der Beurteilung der verschiedenen in Frage kommenden Evaluationsansätze und Strategien, deren Abläufe und aufzuwendenden Budgets.

Prozessevaluation (Process):

Hierbei handelt es sich um die laufende Prüfung und Dokumentierung der Durchführung einer zu evaluierenden Massnahme. Dies gibt den Projektverantwortlichen einerseits Feedback bezüglich der Durchführung einer Massnahme, liefert andererseits aber auch Informationen zur Optimierung.

Produkteevaluation (Product):

Bei der Produkteevaluation steht die Frage nach der Zielerreichung der durchgeführten Massnahme im Zentrum. Die Produktevaluation lässt sich weiter aufteilen in die vier Subkategorien Wirkung (Impact), Effektivität (Effectiveness), Nachhaltigkeit (Sustainability) und Übertragbarkeit (Transportability) (Stuffelbeam, 2002).

Wo Evaluation ansetzt, ist davon abhängig, wie Prozesse gesteuert sind. Dabei lassen sich Input-, Verhaltens- und Output-Steuerung unterscheiden. Die Wahl der Systemsteuerung wirkt sich dabei auf die Durchführungsbedingungen von Evaluationsprojekten aus (Wotawa & Thierau, 1998). Bei einem input-gesteuerten System steht im Zentrum der Evaluation das entsprechende Auswahlverfahren. Dies gilt zum Beispiel bei der Wahl von Professoren an Universitäten in der Schweiz oder etwa in Deutschland. Bei einem verhaltensgesteuerten System konzentrieren sich die Evaluationsmassnahmen dagegen auf die Einhaltung der geltenden Vorgaben. Ein solches System stellt zum Beispiel die öffentliche Verwaltung dar. So ist in einer Bürokratie die Richtigkeit eines Gesetzes durch seinen ordnungsgemässen Beschluss bedingt und nicht durch seinen Nutzen für den Bürger. In output-orientierten Systemen konzentrieren sich Evaluationsvorhaben auf die Ergebnisse, welche ein System liefert. Einzelne Einrichtungen geniessen in heutiger Zeit vermehrte Freiheiten auf organisatorischer Ebene. Dies vor allem im Zug eines von der Wirtschaft entnommenen Steuerungsmodells, welches stark output-orientiert gestaltet ist. Mit zusätzlichen Freiheiten steigen jedoch auch die Verantwortung sowie die Rechenschaftspflicht. Ein solches Konzept stellt damit hohe Anforderungen an die systematische Evaluation solcher Einrichtungen (Böttcher, 2006). Diese Ausrichtung nach dem Output ist vor allem in der Wirtschaft vorherrschend, durchdringt aber immer weitere Bereiche der Gesellschaft. Abb. 13 veranschaulicht die hier erwähnten Zusammenhänge.

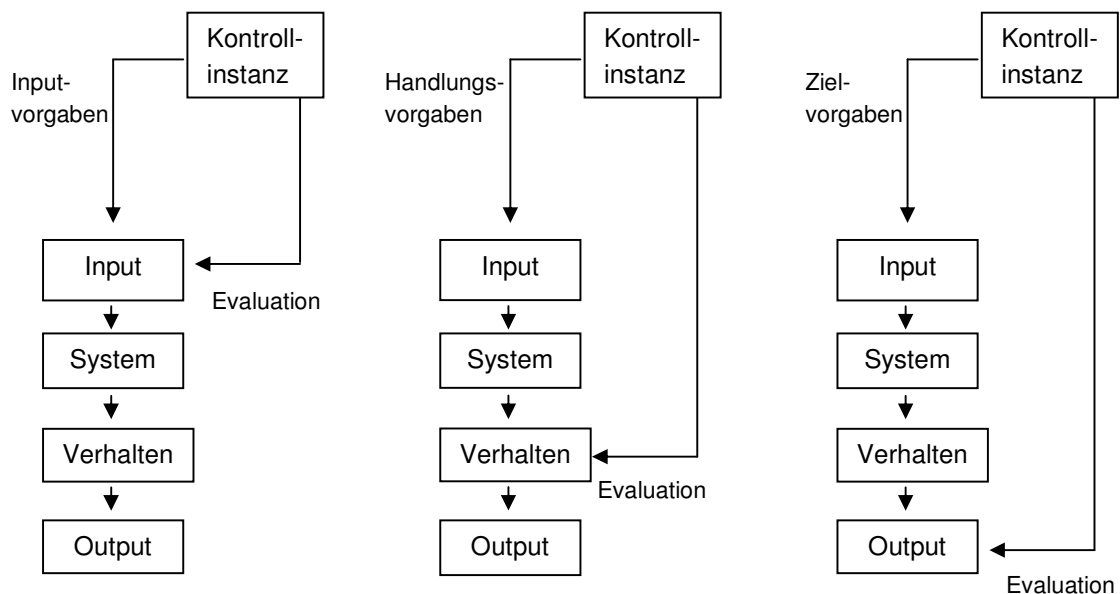


Abb. 13: Ansatzpunkte von Evaluationsvorhaben in Abhängigkeit von der Form der Systemsteuerung (Wottawa & Thierau, 1998).

2.1.12 Evaluationsebenen

Evaluation kann sich auf verschiedenen Ebenen abspielen. Mögliche Ebenen sind das Konzept, das Produkt oder Programm, der Prozess, die Reaktion der Benutzer oder Teilnehmer, der Lernerfolg oder aber auch der Transfer. Ein weit verbreitetes Modell der Evaluationsebenen im Bildungsbereich stellt das vierstufige Modell von Kirkpatrick (1998) dar. Dieses definiert sich über die folgenden Stufen:

- Level 1: Reaction
- Level 2: Learning
- Level 3: Behavior
- Level 4: Results

Bei Level 1 handelt es sich um die Reaktion der Nutzer auf das Produkt oder Programm. Diese äussert sich in der Kundenzufriedenheit. Für Kirkpatrick (1998, S.20) stellt die Kundenzufriedenheit, also die Akzeptanz seitens des Nutzers, eine zentrale Komponente für den Erfolg dar: „Positive reaction may not ensure learning, but negative reaction almost certainly reduces the possibility of its occurring.”

Level 2 umfasst den Umfang, in dem es den Nutzern gelingt, ihr Wissen und ihre Fähigkeiten zu erweitern. Level 3 beinhaltet Änderungen auf der Verhaltensebene, wogegen Level 4 die auf dem Gelernten basierenden Effekte auf die Umwelt meint. Während es

sich bei den Levels 3 und 4, vor allem, um Aspekte des Wissenstransfers handelt, sind die Levels 1 und 2 in der Phase der konkreten Auseinandersetzung mit Produkt oder Programm von Bedeutung.

Das bisher Dargelegte umfasste eine Übersicht über die theoretischen Grundlagen der Evaluation. Wie die gemachten Ausführungen erkennen lassen, handelt es sich bei der Evaluation um keinen einheitlichen Wissensentwurf. Bei näherer Betrachtung zeigt sich die Vielfalt und Breite des Feldes. Da die Evaluation Aspekte verschiedener Wissenschaftstraditionen umfasst, ist die resultierende Heterogenität kaum verwunderlich. Je nach Gegenstand der Evaluation werden daher unterschiedlichste methodische Ansätze relevant. Im Rahmen dieser Arbeit ist die Evaluation im Bereich E-Learning zentral. Daher soll im Folgenden konkret auf Überlegungen zur Evaluation in diesem Gegenstandsbereich eingegangen werden.

2.2 Evaluation im E-Learning

2.2.1 Einleitung

Bei der Evaluation von multimedialen Lernumgebungen besteht das Ziel darin, einzelne Trainingsbausteine sowie deren Kombination im Hinblick auf ein vorher festgelegtes Ziel, beziehungsweise im Hinblick auf Nutzen und Wirkung zu untersuchen. Anhand der Beurteilungsergebnisse soll die Qualität der Weiterbildung gesichert, optimiert oder neuen Anforderungen angepasst werden (Reinmann-Rothmeier et al., 1997). Die oben vorgestellten Evaluationskonzepte können grundsätzlich auch auf den Bereich E-Learning übertragen werden. Trotzdem gilt es, aus der Perspektive der Evaluation einige besondere Aspekte und Eigenheiten zu berücksichtigen.

2.2.2 Annäherung an den E-Learning-Begriff

Unter E-Learning werden sämtliche Formen elektronischer Wissensvermittlung verstanden, bei denen Lerninhalte mit Hilfe von Computern multimedial dargeboten werden. Multimedia meint Computerprogramme, welche sich verschiedener Komponenten (Grafik, Text, Ton und Video) bedienen und diese miteinander kombinieren. Es spielt dabei keine Rolle, ob ein Programm sämtliche erwähnten Medien verwendet oder nur einige. Ein interaktives Multimedia-System erlaubt oder verlangt darüber hinaus die Interaktion mit dem Benutzer (Handke, 2003). Schulmeister (1997, S. 22) definiert Multimedia als „eine interaktive Form des Umgangs mit symbolischem Wissen in einer computergestützten Interaktion“.

Bei E-Learning handelt es sich um eines der Hauptanwendungsgebiete für Multimedia-Systeme. Dieses lässt sich wiederum in die Sub-Kategorien Computer-Based-Training (CBT¹³), Web-Based-Training (WBT¹⁴), E-Learning-Portale¹⁵ und Virtuelle Seminare¹⁶ unterteilen. Multimedia-Systeme lassen sich je nach Art der Auslieferung und Speicherung in die Kategorien Offline-Systeme (CD-ROM¹⁷ oder DVD¹⁸) und Online-Systeme

¹³ Beim Computer-Based-Training befinden sich die Lernprogramme lokal beim Benutzer auf dem Computer gespeichert.

¹⁴ Unter Web-Based-Training wird das Lernen mittels Anwendungen verstanden, welche auf Web-Technologie (Internet und Intranet) basieren.

¹⁵ E-Learning-Portale bieten webbasiert eine Vielzahl an Informationen zu einem bestimmten Gegenstand.

¹⁶ Bei virtuellen Seminaren werden die Lerninhalte von einem Dozenten digital aufbereitet (meist handelt es sich um ein Videobild des Trainers sowie die Visualisierung der Lerninhalte).

¹⁷ CD-ROM ist das Akronym für Compact Disc Read-Only Memory und stellt einen Speicher für digitale Daten dar.

¹⁸ Eine DVD (Digital Video Disc oder Digital Versatile Disc) ist ein digitales Speichermedium.

(Internet) einordnen. Bei Offline-Systemen ist die Nutzung im Rahmen von CBT nur an einem einzelnen CBT-System möglich. Bei Online-Systemen ist dagegen die parallele Nutzung der Inhalte durch mehrere Teilnehmer möglich. Der Einsatz von WBT ermöglicht das kooperative Lernen mehrerer Nutzer (Handke, 2003).

E-Learning kann je nach eingesetzten Technologien sowohl zeitlich und kommunikativ entkoppelt (asynchron, z.B. E-Mail, Weblogs¹⁹, Lernsoftware) als auch gekoppelt (synchron, z.B. Videokonferenzen²⁰, Instant Messaging²¹) erfolgen. Asynchrones E-Learning basiert auf vorgefertigtem Lernmaterial und erfordert einen hohen Grad an Selbstorganisation seitens der Lernenden. Bei synchronem E-Learning besteht dagegen eine direkte Interaktion zwischen Lehrenden und Lernern (Karrasch, 2004). Oft werden beide Formen parallel eingesetzt.

2.2.3 Besonderheiten des E-Learnings

E-Learning zeichnet sich im Gegensatz zu traditionellen Lernformen durch eine Anzahl Besonderheiten aus:

Interaktivität

Die wesentliche Eigenschaft von E-Learning-Systemen ist die Möglichkeit der Interaktion zwischen Benutzer und System. Interaktivität bedeutet, dass der Benutzer nicht bloss Rezipient ist, sondern in den multimedial basierten Lernprozess aktiv eingreift (Baumgartner, 1997). Das Ausmass an Interaktivität ist je nach Programm stark unterschiedlich und kann sich auf einfache Klicks zum Umblättern von Seiten beschränken. Echte Interaktivität jedoch bedeutet, dass der Benutzer nicht nur auf Programmvorgaben reagiert, sondern auch die Kontrolle des Benutzers auf Programmablauf und Programmsteuerung (Handke, 2003). Bei klassischen Print- oder audiovisuellen Medien reduziert sich die Einwirkung auf den Prozessverlauf auf nur äusserliche (indem beispielsweise ein bestimmtes Buch zur Hand genommen, respektive weggelegt wird), oder nur auf einseitige Aktionen (z.B. durch das Überspringen eines Buchkapitels) des linear konzipierten Mediums. Nur schon im Begriff Benutzer spiegelt sich – im Gegensatz zum Hörer oder Seher der klassischen audiovisuellen Medien – ein instrumentelles Verhältnis zum Medium (Baumgartner, 1997).

¹⁹ Ein Weblog ist ein internetbasiertes Tagebuch oder Journal, welches öffentlich zugänglich ist.

²⁰ Eine Videokonferenz ist ein audiovisuelles Kommunikationsverfahren, welches basierend auf unterschiedlichen Technologien Konferenzen in Echtzeit zwischen zwei oder mehreren Teilnehmern ermöglicht.

²¹ Instant Messaging (Nachrichtensofortversand) meint das Kommunizieren in Echtzeit mittels entsprechender Software.

Arnold (2006) nennt weitere Eigenheiten von E-Learning:

Orts- und Zeitflexibilität

E-Learning erfolgt weder orts- noch zeitgebunden. Die grundlegende Unmittelbarkeit von Lernen und Lehren wird durch E-Learning aufgehoben. Lernende wie Lehrende gewinnen dadurch zusätzliche Freiheiten.

Vielfalt an Lerninhalten

Das Internet ermöglicht den schnellen und direkten Zugriff auf Unmengen an Information zu nahezu jedem Gegenstand. Zudem können Experten weltweit mühelos kontaktiert und themenspezifische Weblogs konsultiert werden.

Differenzierung von Lern- und Lehrhandlungen

Durch E-Learning lässt sich der Lernprozess hochgradig individualisieren. Die Bestimmung der Lerngeschwindigkeit sowie die Reihenfolge, in der die Lerninhalte bearbeitet werden, liegen in der Hand des Lernenden. Damit ändert sich auch die Rolle der Lehrenden. Die Betreuung der Lernenden im Lernprozess erhält einen höheren Stellenwert, während die Präsentation und die didaktische Aufbereitung an Gewicht verlieren.

Neue Kooperationsformen

Neben selbstorganisiertem Lernen fördert E-Learning verschiedene Formen von Kooperationen. Als Beispiele hierfür gelten Weblogs oder Online-Communities²². In diesen findet ein oft reger und fruchtbarer Wissensaustausch statt.

Die spezifischen Charakteristika des E-Learnings stellen grundsätzlich potenzielle Stärken und Vorteile gegenüber alternativen Lernformen dar. Der Umgang mit den neuen Medien stellt aber auch hohe Anforderungen und nimmt die Lernenden in die Pflicht: So fordert die Flexibilität von Zeit und Ort ein hohes Mass an Selbstorganisation und Disziplin. Die Vielfalt an Lerninhalten ist nur mit entsprechenden Orientierungs- und Erschließungshilfen erreichbar. Die Differenzierung von Lern- und Lehrhandlung führt auf Seiten der Lerner als auch der Lehrer zu Unsicherheiten und macht den Erwerb zusätzlicher Kompetenzen notwendig. Schliesslich müssen die neuen Kooperationsformen in ein stimmiges E-Learning-Konzept eingebettet sein, sonst drohen sie ungenutzt zu bleiben.

E-Learning verstärkt auch den Paradigmenwechsel, der in der beruflichen Weiterbildung feststellbar ist. Kognitivistische und konstruktivistische Ansätze nehmen den Platz von behavioristisch und instruktionalistisch geprägten didaktischen Ansätzen ein. Die Mög-

²² Eine Online-Community ist eine Gemeinschaft von Personen, welche über das Internet in Kontakt steht.

lichkeiten des E-Learnings führen durch die stark ausgeprägte Bedarfsorientierung und Individualisierung der Lernangebote zu einer Radikalisierung dieser Entwicklung (Ehlers, 2002).

2.2.4 Qualitätssicherung im Bereich E-Learning

Je nach Rahmenbedingungen (Lehrstoff, Personengruppe, Lernumgebung etc.), in die eine Lehrmethode eingebettet ist, fällt deren Effektivität unterschiedlich aus. Nach Fricke (2000) gilt dies auch für den Bereich E-Learning. Er unterscheidet vier Faktoren, welche bei der Konstruktion und Evaluation von E-Learning berücksichtigt werden müssen: die medial gestützte Lern- und Lernumgebung, die Lernvariablen (Lerner), das Lernthema sowie das Lernergebnis. Von diesen Faktoren ausgehend, lassen sich ein präskriptiver und ein deskriptiver Ansatz definieren, welche auch für die Qualitätssicherung von Bedeutung sind. Beim präskriptiven Ansatz sind es die Lernvariablen, das Lernthema und das Lernergebnis, welche die Lernumgebung bestimmen. Dieser Konstruktionsprozess – auch Instruktionsdesign genannt – sollte auf allen Stufen evaluiert werden. Der deskriptive Ansatz beschreibt die umgekehrte Richtung. Hier bestimmt sich das Lernergebnis über die Lernvariablen, die Lernumgebung sowie das Lernthema. Evaluiert wird hierbei in erster Linie die Lernmethode (Lehr- Lernumgebung²³, Lernprogramme etc.). Je nach Perspektive wird damit ein unterschiedlicher Evaluationsansatz relevant.

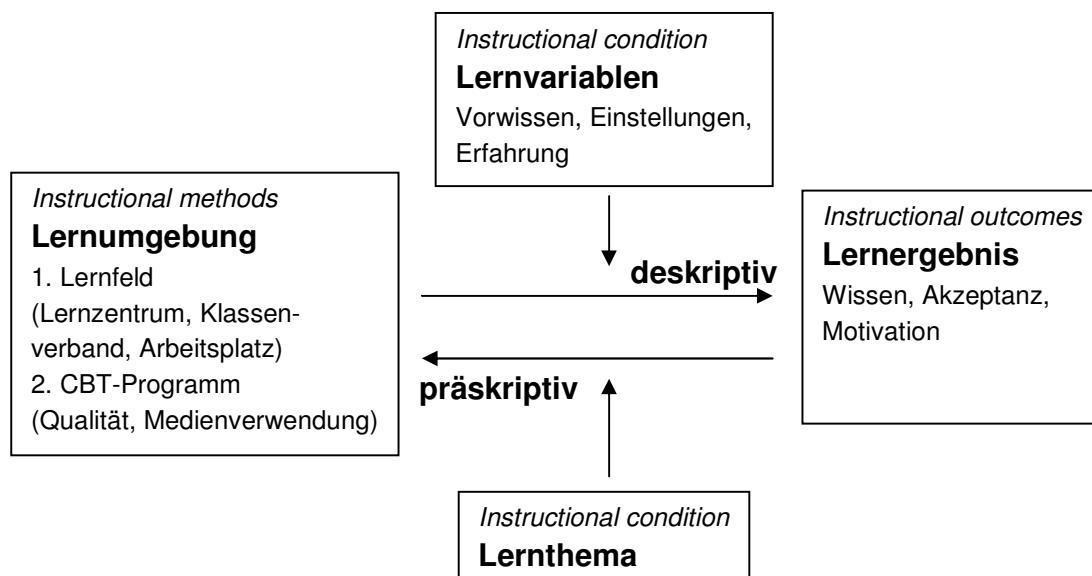


Abb. 14: Paradigma zur Konstruktion und Evaluation multimedialer Lehr- Lernumgebungen (Fricke 1995, zit. nach Fricke, 2000).

²³ Eine Lehr-Lernumgebung beinhaltet sämtliche instrumentellen, zeitlichen, räumlichen und personellen Aspekte einer konkreten Lernsituation.

Da Aus- wie auch Weiterbildung vermehrt auf Basis von E-Learning-Angeboten erfolgen, umfasst die Qualitätsdiskussion in der Weiterbildung verstärkt auch die Frage nach der Qualität im Bereich E-Learning. Dies umso mehr, nachdem die anfängliche Euphorie bezüglich mediengestützten Lernens und Lehrens einer gewissen Ernüchterung Platz gemacht hat (Arnold, 2006). Die Qualitätsdiskussion wird im Bereich E-Learning sehr intensiv geführt. Die Sicherung von Effektivität und Effizienz hat hier eine besondere Bedeutung, da im Fall von E-Learning-Konzepten die Initial- und Folgekosten meist besonders hoch sind (Jäger, 2001).

Je nach Perspektive ergibt sich ein unterschiedliches Qualitätsverständnis. Der produktbezogene Ansatz definiert Qualität als physikalische Eigenschaft. Der anwenderbezogene Ansatz dagegen richtet sich nach den individuellen Präferenzen der Kunden, wobei sich Qualität über die Gebrauchstauglichkeit bestimmt und die Nutzenpräferenz entscheidet. Weiter existiert der fertigungsbezogene Ansatz, welcher auf die Fertigung fokussiert und Standards festlegt, durch deren Einhaltung Qualität entsteht (Ehlers, 2002). Nach Klüber und Löwe (2006) soll man sich bei der Frage danach, was Qualität ausmacht, von Alltagsdefinitionen lösen und sich an professionellen Definitionen orientieren. Ein Beispiel hierfür wäre die ISO 8402, welche Qualität als Gesamtheit von Merkmalen einer Einheit bezüglich ihrer Eignung, festgelegte und vorausgesetzte Erfordernisse zu erfüllen, definiert.

Qualität im E-Learning ist ein vielschichtiges Konzept und umfasst verschiedenste Aspekte. Sie kann sich auf die Einhaltung von Standards, auf technische Anforderungen, auf das Kosten-Nutzen-Verhältnis etc. beziehen. In den meisten Fällen werden unter Qualität im E-Learning jedoch der Kompetenzzuwachs und damit der Lernerfolg verstanden. Dieses Verständnis von Qualität kann als Transformation beschrieben werden und beinhaltet die Weiterentwicklung des Lernenden durch den Lernprozess.

Ein allgemeines Kennzeichen von E-Learning-Projekten ist die heterogene Struktur des Entwicklerteams und die damit verbundene hohe Arbeitsteilung in der Entwicklung. Diese erfordert das Know-how von Softwareentwicklern, Designern, mediendidaktischen Fachleuten wie auch die Mitwirkung von Experten, welche für das Zusammenstellen der Inhalte verantwortlich sind. Jeder dieser Akteure hat nun wiederum eigene Vorstellungen davon, was die Qualität mediengestützten Lernens ausmacht. Weiter lassen sich verschiedene Qualitätsebenen unterscheiden. Arnold (2006) unterscheidet drei verschiedene Ebenen:

- Input-Ebene:

Inputaspekte betreffen die eingesetzten Ressourcen, die Organisation, die allgemeinen Rahmenbedingungen etc. als strukturelle Voraussetzungen.

- **Durchführungsebene:**
Durchführungsaspekte beziehen sich auf das methodisch-didaktische Vorgehen, die Lernberatung, die Lernszenarien sowie die Steuerung des Erstellungsprozesses.
- **Output-Ebene:**
Output-Aspekte beziehen sich auf die erzielten Ergebnisse wie den Kompetenzzuwachs der Teilnehmer, den Prüfungserfolg, die Zufriedenheit, die Akzeptanz etc.

Abb. 15 veranschaulicht die Mehrdimensionalität des Qualitätskonzepts im E-Learning.

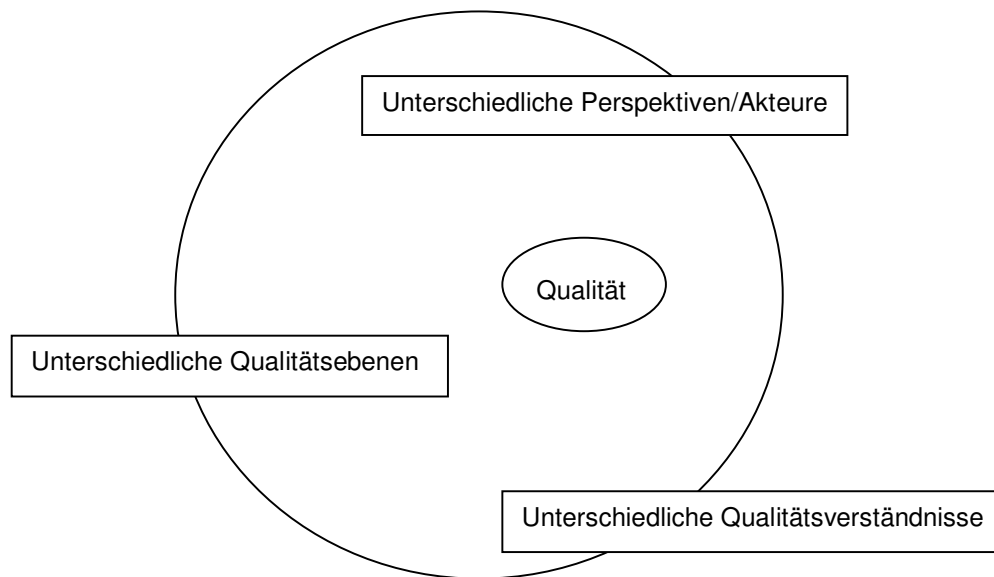


Abb. 15: Qualitätsdimensionen (in Anlehnung an Ehlers, 2002).

Die Qualitätsfrage ist wie hier aufgezeigt sehr komplex. Je nach Akteur und betrachteter Ebene kann die Bedeutung von Qualität sehr unterschiedlich ausfallen. Was unter Qualität letztlich verstanden wird, ist von den konkreten Gegebenheiten im jeweiligen Projekt und den Erwartungen der involvierten Personen bestimmt.

2.2.5 Anforderungen an die Evaluation von E-Learning

Die im Bereich der Bildung entwickelten Evaluationskonzepte lassen sich nicht ohne weiteres auf den Bereich E-Learning anwenden. Zudem führt die spezielle Position der Lerner beim E-Learning dazu, dass sich die Qualitätssicherung stark an diesen orientieren muss.

2.2.5.1 Unterschiede zu traditionellen Lernformen

E-Learning ermöglicht orts- und zeitunabhängiges Lernen. Im Gegensatz zu traditionellen Lernformen, bei denen Lehrpersonen ortsgebunden an Lernende vermitteln, ist die Evaluation damit nicht einfach durch die Beobachtung einer Unterrichtseinheit möglich. Eine Vielzahl der Evaluationsmethoden, die bei traditionellem Lernen in Gruppen Anwendung finden (beispielsweise Beobachtung oder Gruppendiskussion), können aus Kostengründen oder wegen zu hohem Zeitaufwand nicht eingesetzt werden (Baumgartner, 1997). Auf der anderen Seite eröffnet die Digitalisierung der Lehr- und Lernmittel der Evaluation neue Möglichkeiten. So kann beispielsweise dank entsprechender Software mittels Logfiles²⁴ das Verhalten (Navigationsverhalten etc.) der Benutzer auf der Leroberfläche registriert und analysiert werden.

Beim traditionellen Schullernen im Gruppenunterricht sind die demographischen Bedingungen relativ homogen. Auch im Bezug auf die Vorkenntnisse unterscheiden sich die Lernenden meist nur geringfügig. Durch den Lehrplan sind die Lerninhalte sowie die zeitliche Abfolge in der Vermittlung zudem vorgegeben. Beim E-Learning gestaltet sich die Ausgangssituation deutlich anders: Der Zugang erfolgt individuell, was auch zu individuell gestalteten Lernprozessen führt. Inhalt, Zeitpunkt, Intensität und Dauer des Lernprozesses liegen in der Hand der Lerner selber. Medial gestütztes Lernen ist frei von Zugangsvoraussetzungen, jedoch auch in Bezug auf die Zielvorstellungen und die damit verbundenen Motivationsstrukturen keinen Restriktionen unterlegen (Baumgartner, 1997). Beim traditionellen Unterricht ist das Lernziel in Form von Prüfungen, Zeugnissen etc. meist vorgegeben. Beim E-Learning ist dies nicht mehr der Fall. Die verschiedensten Motive bewegen hier die Lerner zur Auseinandersetzung mit den medial dargebotenen Inhalten: die Spannbreite ist hoch und kann von der wenig zielgesteuerten, interessegeleiteten Informationssuche bis zur Absolvierung eines Online-Kurses inklusive Leistungstest reichen. Die Konsequenz hieraus für die Evaluation liegt darin, dass die Anforderungen komplexer werden und neben heterogenen Ausgangssituationen auch individuelle Lernstrategien und Lernziele berücksichtigen müssen. Baumgartner (1997) argumentiert, dass Evaluationen mediengestützten Lernens auf Grund der hohen Komplexität häufig den Einsatz qualitativer Forschungsmethoden verlangen.

2.2.5.2 Orientierung an den Lernenden

Im Gegensatz zur industriellen Produktion herrscht im Bildungsbereich kein einfaches Anbieter-Kunde-Verhältnis. Es handelt sich vielmehr um ein Ko-Produzenten-Verhältnis:

²⁴ Bei Logfiles handelt es sich um die automatische Erfassung aller beim computergestützten Lernen durchgeführten Aktionen, mit dem Ziel der Erhebung vollständiger und differenzierter Daten über das Nutzerverhalten (Tergan, 2000).

E-Learning bietet die Technologie sowie die Inhalte, lernen muss der Nutzer aber selber. Der Lernende wird zum Prosumer (ein Kunstwort aus Producer und Consumer) und ist zu einem grossen Teil selber für die erreichte Qualität im Lernprozess verantwortlich. Qualität wird erst im Prozess des Lernens von den Lernenden selber hergestellt (Arnold, 2006). Die zentrale Rolle der Lerner hat Folgen für die Qualitätsentwicklung im E-Learning. Ehlers (2002) führt hierzu vier Konsequenzen für die Qualitätssicherung von E-Learning auf:

A) Anwenderorientierung ersetzt Technologieorientierung

E-Learning soll technologiegestützt, aber nicht technologieorientiert erfolgen. „Nicht das, was technologisch möglich ist, ist zentral für die Qualität des Angebotes, sondern das, was – mit einer entsprechenden Technologie – inhaltlich, methodisch und situativ auf die Situation des Lernenden hin abgestimmt ist und was den Lernerbedürfnissen entspricht“ (Ehlers, 2002, S. 9). Vermeintlich gute Programme, welche nicht im Sinne der Lerner eingesetzt werden, führen zu schlechteren Lernleistungen als qualitativ weniger hochstehende Programme, die auf die Lerner ausgerichtet sind und deren Ansprüchen und Bedürfnissen entsprechen.

B) Lernerorientierung

Inhalte sollen an den Lernern orientiert und auf deren aktuelle Situation angepasst aufbereitet werden. Über dieses in der Weiterbildung weitherum anerkannte Konzept soll die Entwicklung der Lerner und ihre zukünftige Situation berücksichtigt und die geeignetsten Formen der Vermittlung gefunden werden.

C) Qualität als Produkt des Lernens

Ein E-Learning-Angebot für sich genommen enthält keine Lernqualität, sondern bildet nur den Kontext, in dem Lernen stattfinden kann. Wie oben bereits erläutert, entsteht Qualität erst im Prozess der Nutzung des Lernangebots.

D) Qualitätsförderung

Die Qualität der Lernsoftware ist eine notwendige, aber keine hinreichende Bedingung für erfolgreiches E-Learning. Ausschlaggebend sind in erster Linie die Lernarrangements, die Lernkultur, aber auch die Motivation und Handlungen der Lerner. Lernqualität zeigt sich als ein Konglomerat verschiedenster Komponenten und ist abhängig vom stimmigen Zusammenspiel derselben.

Die hier beschriebene ausserordentliche Bedeutsamkeit der Lernenden für die Qualitätssicherung im E-Learning erfordert zukünftig eine stärkere Ausrichtung der Qualitätsfor-

schung an den Lernern. Die Herausforderung wird sein, die Parameter der Qualität von E-Learning aus Benutzersicht zu bestimmen und die Lernangebote danach auszurichten.

2.2.6 Anforderung an Evaluationsmethoden im E-Learning

Neben der Erfüllung grundlegender Kriterien wissenschaftlichen Arbeitens (Validität, Reliabilität und Objektivität), verlangt die Anwendung von Evaluationsmassnahmen im Bereich E-Learning die Berücksichtigung der Eigenheiten dieses Kontexts. Schott (2000) stellt drei Forderungen an Evaluationsmethoden. Diese sollen adressaten-, evaluator- und theorieorientiert sein:

1) Berücksichtigung der Adressaten

Der Evaluator führt die Evaluation quasi im Namen der Adressaten durch, für welche die Lernumgebung konzipiert ist. Dies bedeutet, dass psychologisch relevante Prozesse der Adressaten bei der Auseinandersetzung mit dem Lernprogramm wie Lernen, Motivation, Emotion und Problemlösen miteinbezogen werden müssen.

2) Ausrichtung an Evaluatoren

Die eingesetzten Methoden sollen den Evaluator dort unterstützen, wo eine Unterstützung wirklich nötig und angebracht ist. Aber die Handhabung der Methoden darf die Evaluatoren kognitiv und emotional nicht überfordern. Oftmals verfügen die Evaluatoren nicht über das fachliche Rüstzeug, um sich mit den komplexen Vorgängen der Lehr- und Lernprozesse differenziert auseinanderzusetzen. Weiter sollen die Methoden leicht in der Handhabung sein.

3) Theoriegeleitete Methoden

Die Methode, die zur Evaluation herangezogen wird, soll theoretisch fundiert sein. „Bei einem theorieorientierten Verfahren ist es möglich, aufgrund von Erfahrungen die Evaluationsmethode gezielt zu verbessern, weil die Struktur des Verfahrens und die Konzepte, die seiner Wirkung zugrunde liegen, offengelegt sind“ (Schott, 2000, S. 112).

Evaluationen sollen weiter effizient und im methodischen Vorgehen ganzheitlich sein sowie neben einer angemessenen Bewertung auch das Kriterium eines vertretbaren Kosten-Nutzen-Verhältnisses erfüllen.

2.2.7 Evaluationsmethoden im E-Learning

Die Qualitätssicherung eines E-Learning-Angebots erfordert die Erhebung konkreter Daten, welche die Grundlage zur Bewertung und Entscheidungsfindung bilden. Die Werkzeuge, die hierfür zum Einsatz kommen, sind vielfältiger Natur. Tergan (2000) schlägt eine Reihe von Methoden vor:

Befragung

Befragungen als reaktive Form der Datenerhebung sind ein wertvolles Hilfsmittel zur Informationsgewinnung und können während des gesamten Evaluationsprozesses eingesetzt werden. In der Planungsphase dienen Befragungen beispielsweise dazu, den Bedarf sowie die gewünschte Ausgestaltung eines E-Learning-Angebots bei der Zielgruppe zu eruiieren. In der Entwicklungsphase kann Befragung im Sinne der formativen Evaluation zur Beurteilung einzelner Aspekte eines Angebots verwendet werden. Besondere Bedeutung erhalten Befragungen in der Einsatzphase als summative Evaluation zur Erhebung der Benutzerakzeptanz.

Im Rahmen des Einsatzes einer Lernumgebung können anhand von Befragungen im Sinne einer summativen Evaluation beispielsweise das allgemeine Nutzungsverhalten der Lerner, der eingeschätzte Nutzen der verschiedenen mediendidaktischen Elemente sowie Verbesserungsvorschläge etc. erfasst werden. Auf diese Weise können Informationen zur generellen Akzeptanz der Lernumgebung gewonnen werden, woraus sich wiederum Möglichkeiten zur Optimierung ableiten lassen.

Beobachtung

Bei der Beobachtung werden sinnlich wahrnehmbare Verhaltensweisen und Ereignisse erhoben. Beobachtungen liefern Ergebnisse, welche diejenigen der Befragung sinnvoll ergänzen oder sogar empirisch unterlegen können. Als konkretes Beispiel kann in diesem Zusammenhang das Usability-Testing (Benutzertests) genannt werden. Im Rahmen der Evaluation eines E-Learningangebots dient diese Methode in der Entwicklungsphase der Beobachtung der Auseinandersetzung eines Benutzers mit dem System. Dadurch können Probleme und Unstimmigkeiten im Interaktionsprozess eruiert werden. Dieses Vorgehen liefert Daten, anhand derer das Lernangebot verbessert und optimiert werden kann. Ein Spezialfall der Beobachtung, welche der technischen Struktur von E-Learning-Systemen entspringt, ist das Verhaltensrecording mittels Logfiles. Damit lässt sich die Interaktion des Benutzers mit dem Lernsystem auf konkreter Ebene verfolgen, was wertvolle Information über die Art und Häufigkeit der Nutzung ergibt.

Tests

Tests bieten die Möglichkeit zur Erhebung quantitativer Daten und ergänzen damit Methoden wie Beobachtung und Befragung, welche qualitative Daten liefern. Während der

Einsatzphase besteht häufig das Ziel, die Wirkung oder den Nutzen (Lernerfolg, Lerntransfer) eines Bildungsangebots zu prüfen und quantitativ zu erfassen. Dies kann anhand von norm- und kriterienorientierten Tests zur Erfassung von Wissen oder Leistung geschehen. So wäre beispielsweise die Untersuchung der Wirkung eines E-Learning-Angebots anhand des Vergleichs von Testresultaten der Nutzer der Lernumgebung mit denjenigen der Besucher einer vergleichbaren Präsenzvorlesung denkbar.

Empirische Untersuchung

Vor allem in der Einsatzphase kann es sinnvoll sein, experimentelle empirische Untersuchungen durchzuführen. Dies zum einen aus strategisch-politischen Gründen, z.B. wenn es darum geht, den Einsatz der finanziellen Mittel, welche in die Projektentwicklung geflossen sind, zu rechtfertigen. Zum anderen kann zum Zweck der Erkenntnisgewinnung eine empirisch-experimentelle Vorgehensweise angestrebt werden. Ein solches Vorgehen wäre beispielsweise angebracht, wenn es darum ginge, die Wirkung verschiedener Elemente einer Lernumgebung auf ihre Wirkung hin zu prüfen. Hierzu könnte eine experimentelle Studie zur Erfassung der Effekte dieser Lernmaterialien durchgeführt werden.

Expertenevaluation

Zusätzlich zu diesen empirisch ausgerichteten Verfahren, existieren noch die Evaluationsmethoden, welche von Experten durchgeführt werden. Expertenevaluationen haben bestimmte Vorteile: sie sind kostengünstig, effizient, führen unmittelbar zu konkreten Ergebnissen, sind kurzfristig plan- und schnell durchführbar. In der folgenden Tabelle sind einige dieser Methoden aufgeführt.

Tab. 4: Expertenmethoden der Evaluation (in Anlehnung an Schenkel, 2002).

| Methode | Beschreibung |
|-------------------------|--|
| Screening | Beurteilung einer Anzahl von Bildschirmseiten nach vorgegebenen Kriterien. |
| Cognitive Walkthrough | Lösen prototypischer Aufgaben zur Aufdeckung möglicher Probleme. |
| Heuristische Evaluation | Prüfung einzelner Dialogelemente nach Usability-Kriterien. |
| Cooperative Walkthrough | Prüfung einer Software durch ein Team von Designern, Entwicklern etc. |
| Eigenschaftsinspektion | Beurteilung der relevanten und benötigten Eigenschaften einer Software. |

Im Bereich E-Learning erfreuen sich weiter Beurteilungen anhand von Kriterienkatalogen grosser Beliebtheit. Dabei wird ein Bildungsangebot von einem Experten anhand eines

Kriterienkatalogs umfassend²⁵ eingeschätzt. Diese vergleichsweise einfach durchzuführende, zeit- und kostengünstige Methode berücksichtigt jedoch weder den konkreten Anwendungskontext in der Lernsituation noch den Lernprozess selber (Arnold, 2006). Weitere Probleme ergeben sich durch die oft unklare Zuteilung der Produkte zur Programmtypologie, den grossen Interpretationsspielraum in der Anwendung der Kriterien, der je nach Programmtyp unverständlichen oder ungeeigneten Kriterien sowie der unterschiedlichen Auslegung der Kriterien je nach Programmtyp (Biffi, 2002). Obwohl solche Expertenratings wertvolle Hinweise über die Qualität eines Produkts geben können, greifen sie daher oft zu kurz.

Mit den Ausführungen zu den Grundlagen der Evaluation sowie deren Anwendung auf den Gegenstand E-Learning ist die theoretische Basis gelegt worden, anhand derer die Lernumgebung PTO theoriegeleitet und systematisch evaluiert werden kann.

²⁵ Der Kriterienkatalog der Spezifikation PAS 1032-1/2 (Publicly Available Specification), welcher sich nach der ISO-Norm 9241 richtet, umfasst beispielsweise die sieben Bereiche Rahmenbedingungen, technische Aspekte, Datenspeicherung und Verarbeitung, Funktionalitäten, (lern)theoretische Aspekte, Informationskodierung und Formate/Gestaltung.

2.3 Literaturverzeichnis

- Abs, H. J., Maag-Merki, K. & Klieme, E. (2006). Grundlegende Gütekriterien für Schulevaluationen. In W. Böttcher, H. G. Holtappels & M. Brohm (Hrsg.), *Evaluation im Bildungswesen* (S. 97-108). Weinheim: Juventa.
- American Evaluation Association (AEA), Task Force on Guiding Principles for Evaluators (1995). Guiding Principles for Evaluators. *New Directions for Program Evaluation*, 66, 19-26.
- Arnold, P. (2006). Qualitätsentwicklung im E-Learning – Ansätze, Herausforderungen und Perspektiven. In B. Schwarz & D. Behrmann (Hrsg.), *Integratives Qualitätsmanagement – Perspektiven und Praxis der Organisations- und Qualitätsentwicklung in der Weiterbildung* (S. 79-114). Bielefeld: Bertelsmann.
- Baumgartner, P. (1997). Evaluation vernetzten Lernens: 4 Thesen. In H. Simon (Hrsg.), *Virtueller Campus. Forschung und Entwicklung für neues Lehren und Lernen* (S. 131-146). Münster: Waxmann.
- Baumgartner, P. (2005). Corporate E-Learning bewerten. In K. Hilbers (Hrsg.), *Stolpersteine beim Corporate E-Learning* (S. 97-122). München: Oldenbourg.
- Behrmann, D. (2006). Qualitätsmanagement ist nicht alles – betrifft aber vieles. Zur systematischen Verortung eines Gestaltungskonzepts für die Weiterbildungsorganisation. In B. Schwarz & D. Behrmann (Hrsg.), *Integratives Qualitätsmanagement – Perspektiven und Praxis der Organisations- und Qualitätsentwicklung in der Weiterbildung* (S. 21-77). Bielefeld: Bertelsmann.
- Beywl, W. (1988). *Zur Weiterentwicklung der Evaluationsmethodologie*. Frankfurt am Main: Lang.
- Biffi, C. (2002). *Evaluation von Bildungssoftware* [On-line]. Available: <http://www.medienpaed.com/02-1/biffi1.pdf>
- Bortz, J. & Döring, N. (2002). *Forschungsmethoden und Evaluation*. Berlin: Springer.
- Böttcher, W. (2006). Bildungsstandards und Evaluation im Paradigma der Outputsteuerung. In W. Böttcher, H. G. Holtappels & M. Brohm (Hrsg.), *Evaluation im Bildungswesen* (S. 39-49). Weinheim: Juventa.
- Ehlers, U. (2002). *Qualität im E-Learning: Der Lernende als Grundkategorie bei der Qualitätssicherung* [On-line]. Available: <http://www.medienpaed.com/02-1/ehlers1.pdf>
- Fricke, R. (2000). Qualitätsbeurteilung durch Kriterienkataloge. In P. Schenkel, S. O. Tergan & A. Lottmann (Hrsg.), *Qualitätsbeurteilung multimedialer Lern- und Informationssysteme* (S. 75-88). Nürnberg: BW Bildung und Wissen.
- Fricke, R. (2002). Evaluation von Multimedia. In L. J. Issing & P. Klimsa (Hrsg.), *Information und Lernen mit Multimedia und Internet* (3. vollst. überarb. Aufl.) (S. 445-463). Weinheim: Beltz.
- Guba, E. G. & Lincoln, Y.S. (1989). *Fourth Generation Evaluation*. Newbury Park: Sage.
- Handke, J. (2003). *Multimedia im Internet*. München: Oldenbourg.
- Jäger, W. (2001). E-Learning. *Personal - Zeitschrift für Human Resource Management*, 53 (7), 374-379.

- Joint Committee on Standards for Educational Evaluation (1994). *The Program Evaluation Standard* (2nd ed.). Newbury Park: Sage.
- Karrasch, H. (2004). Fortbildung Online. In H. Karrasch, C. Kraugartner & B. Prowaznik (Hrsg.), *E-Learning – Wunschtraum oder Realität?* (S. 25-48). Wien: Börsedruck.
- Kirkpatrick, D. L. (1998). *Evaluating Training Programs* (2nd ed.). San Francisco: Berrett-Koehler.
- Klüber, K. & Löwe, C. R. (2006) *Qualitätsmanagement und Zertifizierung in Bildungsorganisationen*. Augsburg: Ziel.
- Madaus, E. & Stufflebeam D. L. (2000). Program Evaluation: a Historical Overview. In D. L. Stufflebeam, G. F. Madaus & T. Kellaghan (Eds.), *Evaluation Models: Viewpoints on Educational and Human Services Evaluation* (2nd ed.), (S. 3-18). Norwell: Kluwer.
- Mertens, D. & McLaughlin, J. A. (2004). *Research and Evaluation Methods in Special Education*. Thousand Oaks: Sage.
- Meyer, W. & Höhns, G. (2002). *Was ist Evaluation?* Bundesinstitut für Berufsbildung: Bonn.
- Reinmann-Rothmeier, G., Mandl, H. & Prenzel, M. (1997). Qualitätssicherung bei Multimedialen Lernumgebungen. In H. F. Friedrich, G. Eigler, H. Mandl, W. Schnotz, F. Schott & N. M. Seel (Hrsg.), *Multimediale Lernumgebungen in der betrieblichen Weiterbildung* (S. 267-332). Neuwied: Luchterhand.
- Rindermann, H. (2001). *Lehrevaluation – Einführung und Überblick zu Forschung und Praxis der Lehrveranstaltungsevaluation an Hochschulen*. Landau: Empirische Pädagogik.
- Rindermann, H. (2003). Lehrevaluation an Hochschulen: Schlussfolgerungen aus Forschung und Anwendung für Hochschulunterricht und seine Evaluation. *Zeitschrift für Evaluation*, 3 (2), 233-256.
- Rossi, P. H., Lipsey, M. W. & Freeman, H. (2004). *Evaluation. A Systematic Approach*. Thousand Oaks: Sage.
- Rowntree, D. (1992). *Exploring Open and Distance Learning*. London: Kogan Page.
- Schenkel, P. (2000). Ebenen und Prozesse der Evaluation. In P. Schenkel, S.O. Tergan & A. Lottmann (Hrsg.), *Qualitätsbeurteilung multimedialer Lern- und Informationssysteme* (S. 52-74). Nürnberg: BW Bildung und Wissen.
- Schott, F. (2000). Evaluation aus theoriegeleiteter, ganzheitlicher Sicht. In P. Schenkel, S.O. Tergan & A. Lottmann (Hrsg.), *Qualitätsbeurteilung multimedialer Lern- und Informationssysteme* (S. 106-124). Nürnberg: BW Bildung und Wissen.
- Schulmeister, R. (1997). *Grundlagen hypermedialer Lernsysteme* (2. Aufl.). München: Oldenbourg.
- Scriven, M. (1991). *Evaluation Thesaurus* (4th ed.). Newbury Park: Sage.
- Stockmann, R. (2006). Qualitätsmanagement und Evaluation im Vergleich. In W. Böttcher, H. G. Holtappels & M. Brohm (Hrsg.), *Evaluation im Bildungswesen* (S. 23-38). Weinheim: Juventa.
- Stufflebeam, D. L. (2000). The CIPP Model for Evaluation. In D. L. Stufflebeam, G. F. Madaus & T. Kellaghan (Eds.), *Evaluation Models: Viewpoints on Educational and Human Services Evaluation* (2nd ed.), (S. 279-317). Norwell: Kluwer.

Stufflebeam, D. L. (2002). *CIPP Evaluation Model Checklist* [On-line]. Available: <http://www.wmich.edu/evalctr/checklists/cippchecklist.pdf>

Stufflebeam, D. L. (2007). *Evaluation Theory, Models, and Applications*. San Francisco: Wiley.

Tergan, S.O. (2000). Grundlagen der Evaluation. In P. Schenkel, S.O. Tergan & A. Lottmann (Hrsg.), *Qualitätsbeurteilung multimedialer Lern- und Informationssysteme* (S. 22–51). Nürnberg: BW Bildung und Wissen.

Vedung, E. (1999). *Evaluation im öffentlichen Sektor*. Wien: Böhlau.

Wottawa, H. & Thierau, H. (1998). *Evaluation* (2. Aufl.). Bern: Huber.

3.

Evaluationskonzept

Abstract

Die folgenden Ausführungen setzen sich mit Adaptivitätsmassnahmen und deren Evaluation im E-Learning am Beispiel der webbasierten Lernumgebung „Psychopathology Taught Online“ (PTO) auseinander. PTO soll ergänzend zum universitären Lehrangebot als digitales Curriculum eingesetzt werden. Es zeichnet sich dadurch aus, dass Wissen inhaltlich adaptiv vermittelt wird. Dem Lerner werden dazu konkrete, an seinen momentanen Wissensstand angepasste Lernempfehlungen gegeben, was den Lernprozess individualisierter und effizienter gestaltet. In seiner Gestaltung baut PTO auf kognitionspsychologischen Theorien zu Assimilation und Akkomodation beim Lernen auf. Diese Theorie ist entwicklungspsychologisch gut hergeleitet, in adaptiven Verfahren jedoch noch nicht in der Praxis evaluiert worden. Während der Entwicklungs- und Einsatzphase der neuen Lernumgebung werden deswegen verschiedene evaluative Massnahmen ergriffen, die der Qualitätssicherung und der Qualitätskontrolle dienen. Zentral ist dabei die Evaluation der Adaptivität, da diese einen projektübergreifenden Aspekt von allgemeinerer lernpsychologischer Bedeutung darstellt.

3.1 Einleitung

In den letzten Jahren hat sich der Einsatz von E-Learning an den Universitäten massiv verstärkt. Vielfältige Angebote ermöglichen es den Studenten, im Selbststudium den in den Vorlesungen behandelten Stoff zu vertiefen. Die Vorteile liegen dabei vor allem in der Individualisierung und der Selbstorganisation des Lernens. Die Lernumgebungen, die dabei zum Einsatz kommen, müssen aber bestimmten Kriterien gerecht werden, damit sich ihr Einsatz rechtfertigt. Es sind daher Massnahmen nötig, die helfen, diesen Qualitätslevel zu erreichen. Der aktuelle Entwicklungsstand dient dabei als Grundlage, um die entsprechenden Massnahmen zur Verbesserung treffen zu können. Diesen Entwicklungsstand mit den geeigneten Mitteln zu erfassen, ist Aufgabe der Evaluation. Sie sollte während des gesamten Entwicklungsprozesses immer wieder die Systemkomponenten überprüfen. Dabei wird auf ein umfangreiches Methodenrepertoire zurückgegriffen. Dieser Ansatz wird auch bei der Evaluation von PTO verfolgt. Einerseits ist dabei das Ziel, die Lernumgebung möglichst optimal den Bedürfnissen der Benutzer anzugleichen, andererseits soll auch der effektive Lernerfolg, der mit PTO erzielt werden kann, evaluiert werden. Da PTO auf eine adaptive Wissensvermittlung abzielt, ist die Überprüfung derselben ein zentraler Aspekt.

3.2 Chancen der Adaptivität im E-Learning

Aus konstruktivistischer Perspektive repräsentiert sich das Wissen eines Menschen über die Welt in kognitiven Strukturen. Die einzelnen Wissens Elemente stehen nicht diskret für sich, sondern sind in einer hierarchischen Struktur organisiert, aufeinander bezogen und bilden eine Ordnung. Neues Wissen wird also in bereits bestehende Schemata eingefügt. Piaget (1976) bezeichnete diesen Prozess mit Assimilation. Reize und Informationen aus der Umwelt werden dabei aufgenommen, verarbeitet und in die bestehenden Strukturen eingegliedert.

Kommt neues Wissen hinzu, hat das jedoch Auswirkung auf das gesamte Strukturgefüge. Die Modifikation und Anpassung einer Wissensstruktur durch ein neu gelerntes Element wird von Piaget Akkomodation genannt. Assimilation und Akkomodation werden von ihm als Teilprozesse bei der Bildung eines kognitiven Gleichgewichts im Wissenserwerb verstanden. Ausgehend von diesen Ideen, folgert Ausubel (1980/81, 2000), dass sich Lernen dann am wirkungsvollsten gestaltet, wenn neue Elemente in eine bereits fest verankerte, stabile Struktur integriert werden. Aus pädagogischer Sicht muss das Ziel sein, Wissensinhalte so zu vermitteln, dass diese in Relation zu bereits bestehendem Wissen gelernt werden und sich dadurch optimal in bestehende Schemata einfügen können. Dadurch ergeben sich stabile aufeinander bezogene Strukturen. Das Wissen kann in einem weiteren Schritt vertieft und progressiv differenziert werden. Bevor mit dem Lernen be-

gonnen wird, gilt es den aktuellen Wissenstand (Vorwissen) zu diagnostizieren und damit die bestehende Struktur zu erfassen. Dadurch kann fehlendes oder falsches Wissen detektiert und der Lernprozess darauf ausgerichtet werden. Ziel ist es, das Vorwissen zu aktivieren und an dieses anzuknüpfen. Lernen erfolgt somit nicht planlos, sondern adaptiv, da es sich an bestehendem Wissen orientiert. Der primäre Vorteil der Adaptivität liegt darin, dass durch eine Statusdiagnose die Basis geschaffen wird, um dem Lerner Lernempfehlungen geben und ihm diejenigen Wissensinhalte präsentieren zu können, über die er noch nicht verfügt respektive von denen er inadäquate Vorstellungen aufgebaut hat. Dieses Vorgehen verspricht demnach den höchsten Lernerfolg (Leutner, 1992). Desweiteren wird unnötiges Doppellernen vermieden und so ein Effizienzgewinn im Lernprozess erzielt. Diese Vorteile kommen besonders beim E-Learning zum Tragen, da hier der erzielbare Grad der Individualisierung und Selbstorganisation des Lernprozesses durch die technischen Möglichkeiten einer individuellen Zusammenstellung von Materialien sehr hoch ist.

3.3 Das Fallbeispiel: Die E-Learning-Umgebung PTO

Die Entwicklung der E-Learning-Plattform „Psychopathology Taught Online“ (PTO)²⁶ ist ein Projekt, das vom Swiss Virtual Campus (SVC)²⁷ unterstützt wird. PTO soll im Studiengang Psychopathologie des Erwachsenenalters zum Einsatz kommen und den Lernenden einen multimedialen Zugang zu Lerninhalten der Psychopathologie bieten. Daneben sind noch weitere Einsatzmöglichkeiten im Rahmen der klinischen Psychologie, der Psychiatrie oder der postgradualen Weiterbildung denkbar. Das Konzept von PTO basiert auf dem Ansatz des blended learning und soll damit ergänzend zu weiteren universitären Angeboten eingesetzt werden (Streule, Egli, Oberholzer & Läge, 2005). Zusätzlich zu qualitativ hochstehend aufbereiteten Lerninhalten zu psychopathologischen Störungsbildern besteht das Ziel des Projekts PTO darin, technische Verfahren zur Realisierung eines Adaptiven Tutoriellen Systems (ATS) zu entwickeln. Durch ein solches System soll sichergestellt werden, dass sich der Lernprozess der PTO-Nutzer möglichst effizient gestaltet. Um diese Adaptivität gewährleisten zu können, wird ein ähnlichkeitsbasiertes Verfahren zur Diagnostik von Wissen angewendet (Läge & Daub 1998; Läge, 2001): Das diagnostizierte Wissen wird auf der Basis eines Strukturvergleiches mit Expertenwissen verglichen. Diejenigen Störungsbilder, die im Vergleich zum Experten die grössten Unterschiede zeigen, werden erneut zum Lernen empfohlen (für eine ausführlichere methodische Darstellung der Funktionsweise eines Strukturvergleichs siehe Streule et al., 2005).

²⁶ <http://www.pto.unizh.ch>

²⁷ <http://www.virtualcampus.ch>

PTO bietet somit einerseits multimedial aufbereitete Lerninhalte, mittels derer sich der Lerner ergänzend zur Vorlesung im Selbststudium vertieft mit den Störungsbildern der Psychopathologie auseinandersetzen kann. Andererseits zielt PTO auf individuell angepasstes Lernen ab, welches einen effizienten Lernprozess garantieren soll. Im Folgenden soll kurz auf zentrale Aspekte von PTO eingegangen werden.

3.3.1 Curricula

Der Lernprozess innerhalb von PTO gliedert sich in zwei Curricula, die der Lerner sequentiell durchläuft. Im ersten Curriculum wird in 20 multimedia-basierten Lektionen eine repräsentative Auswahl von 20 Störungsbildern gelernt. Hierbei geht es vor allem darum, dem Lerner einen Überblick über ein breites Spektrum des Faches zu vermitteln. Die 20 Störungsbilder sollen nach dem Lernen eine stabile und vor allem korrekte kognitive Grundstruktur bilden (assimilatives und akkomodatives Lernen), an die später im Curriculum 2 weitere Elemente assimilativ angehängt werden können. Begleitend zum ersten Curriculum erfolgt eine Wissensdiagnose: Anhand von Ähnlichkeitsurteilen, die der Lerner zwischen allen Paaren einzelner Störungsbilder abgibt, lassen sich Wissenskarten berechnen. Eine solche Karte, in der jedes Störungsbild eine bestimmte Position einnimmt, ist vergleichbar mit einer Landkarte: Störungsbilder, die nahe beieinander liegen, sind einander ähnlicher als solche, die weiter entfernt voneinander positioniert sind. So lässt sich die individuelle kognitive Struktur eines Wissensbereichs abbilden. Durch Vergleich einer Lernerkarte mit einer Expertenkarte zeigt sich somit mangelhaftes Wissen (dieses Verfahren der Prokrustes-Transformation ist bei Streule et al., 2005 beschrieben). Ist die Diskrepanz zwischen Lerner- und Expertenwissen für einzelne Störungsbilder zu hoch, werden für diese Störungsbilder spezifische Lernempfehlungen gegeben (neben Repetitionsempfehlungen vor allem einzelne Übungen, die sich mit den Relationen zwischen Störungsbildern beschäftigen). Curriculum 1 endet, wenn der akkomodative Prozess der Ausbildung einer korrekten kognitiven Karte abgeschlossen ist.

Das zweite Curriculum dient dem Lernen von insgesamt 32 zusätzlichen Störungsbildern. Diese sollen sich über die oben beschriebenen Prozesse assimilativ in die bereits vorhandene Struktur einfügen. Auch in diesem Fall wird durch kontinuierliche Diagnostik sichergestellt, dass neu gelerntes Wissen richtig verarbeitet und in bestehende Strukturen integriert wird.

Daneben bietet PTO ausgewählte Themen der Psychopathologie. Dies umfasst eine Einführung in den Fachbereich sowie spezifische, relevante Aspekte der Psychopathologie wie therapeutische Aspekte oder verursachende Faktoren, welche sonst in PTO ausgeklammert bleiben (sofern sie keinen Einfluss auf die Klassifikation haben).

3.3.2 Lektionen

PTO orientiert sich inhaltlich am ICD-10, dem hierarchisch organisierten Klassifikationsschema psychischer Störungen, auf das sich die WHO (World Health Organization) beruft²⁸. Die Lernplattform ist hierarchisch nach den einzelnen Kategorien geordnet. Diese sind wiederum in einzelne Lektionen gegliedert, wobei sich eine Lektion mit einem spezifischen Störungsbild befasst (z.B. die Lektion bipolare Störung, die der Kategorie affektive Störungen zugeordnet ist). Die Elemente der einzelnen Lektionen sowie deren Anordnung bleiben über alle Lektionen identisch. Zu Beginn einer Lektion wird der Lerner anhand eines konkreten Beispiels in die Thematik eingeführt. In einem nächsten Punkt werden die Lernziele der Lektion definiert. Weiter werden allgemeine Informationen zum Störungsbild sowie zum Störungsverlauf, zur Verbreitung etc. vermittelt. Kernpunkt einer jeden Lektion sind die ICD-Kriterien, die das Erscheinungsbild einer Störung beschreiben. Anhand eines Fallbeispiels kann das Gelernte praxisorientiert verankert werden. Am Schluss einer jeden Lektion hat der Lerner überdies die Möglichkeit, sein Wissen im Selbsttest zu überprüfen und sich über das Erreichen der Lernziele zu vergewissern. Zudem wird eine Literaturliste präsentiert und mit externen Hyperlinks auf Webseiten verwiesen, auf denen der vermittelte Stoff weiter vertieft werden kann.

3.3.3 Adaptivitätsmassnahmen

PTO zielt auf eine adaptive Wissensvermittlung ab. Wie oben beschrieben, findet während den Curricula 1 und 2 jeweils begleitend eine Wissensdiagnose statt, auf Grund derer individuelle Lernempfehlungen abgegeben werden. Das Feedback, das der Lerner erhält, umfasst drei Massnahmen: Wird bei der Wissensdiagnose das Lernziel (genügende Übereinstimmung der Positionen eines Störungsbildes in der individuellen Karte und der Expertenkarte) nicht erreicht, werden dem Lerner die am weitesten abweichenden vier Störungsbilder zur nochmaligen Bearbeitung empfohlen. Das mangelhafte Wissen soll durch die Repetition korrigiert werden. Zudem werden dem Lerner (als zweite Massnahme) Übungen gegeben. Diese sind so aufgebaut, dass Gemeinsamkeiten von beziehungsweise Unterschiede zwischen einem Paar von Störungsbildern evident werden, von denen genau eines in seinem Wissen relational falsch verankert ist. Dies soll dazu dienen, die strukturelle Organisation des bislang falsch verankerten Störungsbildes durch Vergleich mit einem korrekt verankerten zu verbessern. Erreicht der Lerner nach wiederholten Übungen in der Wissensdiagnose das Lernziel immer noch nicht, so wird ihm (gewissermassen als „letzte Massnahme“) demonstrativ mit Hilfe eines geometrischen Vergleichs seiner eigenen kognitiven Karte und der Expertenkarte gezeigt,

28 Die kategoriale Struktur wird kritisiert von Egli, Schlatter, Streule & Läge (2006).

wo die mangelhaft gewussten Störungsbilder innerhalb der Struktur zu liegen kommen sollten. Am Ende des Prozesses soll ja neben einem korrekten Faktenwissen auch eine korrekte relationale Verankerung der einzelnen Wissenskomponenten zueinander bestehen.

3.4 Evaluationsplan

Für die Evaluation sind vor allem zwei Aspekte von Bedeutung: Erstens das Produkt in Form der Lernumgebung. Zweitens steht die Verwendung desselben sowie die dadurch erzielten Effekte (Lernen und Lernerfolg) im Fokus. Um die Qualität der Lernumgebung sicherzustellen, ist eine umfassende Evaluation notwendig, die die Basis für Entscheidungen im Entwicklungsprozess bildet. Für eine web-basierte Lernumgebung ist es beispielsweise besonders wichtig, dass sie den Usability-Standards genügt, da viele potentielle Lerner noch wenig mit E-Learning vertraut sind und ein funktionierendes Tool dazu beiträgt, Berührungspunkte abzubauen und die Dropout-Quote zu reduzieren. Eine Evaluation klärt weiter die Frage, ob PTO den gestellten Ansprüchen der verschiedenen Stakeholder genügt und dient damit auch nicht zuletzt als strategisch-politische Entscheidungsgrundlage.

Die Lernumgebung PTO befindet sich zum Zeitpunkt der Abfassung dieses Beitrags noch in der Produktionsphase. Die Evaluation von PTO wird durch eine Person vorgenommen, die nicht direkt an der Produktion beteiligt ist. Damit soll genügend Distanz und Unabhängigkeit zu den Inhalten und Verfahrensschritten gewahrt werden, um die Evaluation mit ausreichender Objektivität vornehmen zu können.

In der Produktionsphase liegt das Hauptaugenmerk der Evaluation aller fertig gestellten Lektionen und Verfahrensschritte auf den software-ergonomischen Aspekten mit dem Ziel der Optimierung der Usability (Benutzungsfreundlichkeit) der Lernumgebung. Diese Evaluation dient in erster Linie der Qualitätssicherung (formative Evaluation).

Während der Einsatzphase besteht die Aufgabe der Evaluation vor allem darin zu prüfen, ob mittels PTO der gewünschte Lernerfolg tatsächlich erzielt werden kann und ob der verwendete Adaptivitätsmechanismus tatsächlich greift (summative Evaluation; Merstens, 2005). Bezüglich der Entwicklung von PTO ergibt sich somit aus einer umfassenderen Perspektive eine dreifache Aufgabenstellung für die Evaluation: Erstens soll die Lernumgebung anhand verschiedener Methoden (Expert Review, Usability-Test, etc.) evaluiert werden. Damit soll die Grundlage zur Verbesserung der Benutzungsfreundlichkeit und der Benutzerakzeptanz geschaffen werden. Zweitens soll das Funktionieren der Adaptivitätsmassnahmen untersucht werden. Drittens ist die Effektivität von PTO in einer Wirkungsanalyse zu prüfen. Diese umfasst zum einen den Lernerfolg, der durch die Struktur und Gestaltung der Lernumgebung ermöglicht wird und der den Einsatz der Lernumgebung schlussendlich legitimiert. Andererseits soll hier

die Akzeptanz bei den Stakeholdern (vor allem bei Studierenden, die damit lernen, und Dozenten, die PTO einsetzen) erfasst werden.

3.4.1 Evaluation auf Lektionenebene

Für die Evaluation einer webseitenbasierten Lernumgebung bieten sich verschiedene Massnahmen an. Um der Usability Rechnung zu tragen, wird PTO nach software-ergonomischen Kriterien überprüft. Dabei werden mit der Expert Review und dem Usability-Labortest zwei Methoden verwendet, die sich in der Praxis besonders bewährt haben (Schweibenz & Thissen, 2003). Diese Verfahren dienen in erster Linie dem Aufdecken von Usability-Problemen und bilden damit die Grundlage zur Optimierung der Oberflächengestaltung von PTO sowie der Interaktion zwischen Lerner und Lernumgebung.

Die Expert Review gehört zu den expertenorientierten Methoden und eignet sich besonders in der Frühphase der Entwicklung, um erste Usability-Fehler auszumerzen. Wie der Name schon sagt, wird dabei die Benutzeroberfläche vom Evaluator nach software-ergonomischen Kriterien beurteilt. Dieser orientiert sich bei der Beurteilung an Richtlinien für benutzerfreundliche Gestaltung. Die Resultate können im Team diskutiert, entsprechende Änderungen beschlossen und diese sogleich implementiert werden. So werden die Voraussetzungen geschaffen, um im Labor-Test fundamentale Usability-Probleme zu eruieren, die sich meist erst in der Interaktion zwischen dem eigentlichen Benutzer und dem System zeigen.

Im Usability-Labortest wird die didaktische Gestaltung von PTO von potentiellen Benutzern in einer Testsituation anhand von konkreten Aufgaben geprüft (Nielsen, 1993). Z.B. bestand eine Aufgabe darin, innerhalb eines Moduls spezifische Inhalte zu finden. Die Aufgaben sollen von den Testpersonen anhand der auf der Lernumgebung angebotenen Funktionen bewältigt werden, wobei vor allem Usability-Probleme sichtbar werden, die sich aus dem Arbeitsablauf ergeben. Der Labor-Test bietet damit einen Einblick in die tatsächliche Handhabung des Systems und liefert Informationen zur weiteren Verbesserung der Usability.

Während eines ersten geplanten Testlaufs, bei dem PTO parallel zu einer Grundvorlesung in Psychopathologie zum Einsatz kommt, wird auf der Lernumgebung ein vom PTO-Projektteam moderiertes Forum eingerichtet. Dadurch soll den PTO-Benutzern die Möglichkeit gegeben werden, offenes Feedback und Anregungen zu den unterschiedlichsten Aspekten der Lernumgebung zu geben.

3.4.2 Evaluation der Adaptivitätsmassnahmen

Um die Evaluationsmassnahmen im Bereich der Adaptivität verständlich zu machen, muss kurz auf die Funktionsweise der Wissensdiagnose eingegangen werden. Diese wird

über paarweise Ähnlichkeitsvergleiche zwischen den einzelnen Störungsbildern erhoben. Mit Nonmetrischer Multidimensionaler Skalierung (NMDS) lassen sich diese Ähnlichkeitswerte als Distanzwerte verrechnen, in ein Raummodell überführen und als kognitive Karte veranschaulichen (Borg & Groenen, 1997; Läge, 2001). Diese Karte zeigt die kognitive Struktur für den Wissensbereich psychopathologischer Störungsbilder des Lernalers. Seine Karte kann nun mit der eines Experten verglichen werden, wobei über Diskrepanzen zwischen den beiden Karten mangelhaftes Wissen detektiert werden kann (Streule et al., 2003). Da nach der Lernempfehlung jeweils erneut eine Wissensdiagnose stattfindet, kann über Veränderungen in der Kognitiven Karte auf Veränderungen des Wissens geschlossen werden. Das Lernen zeigt sich also in der Karte und die Veränderung der Karte wird damit als Grundlage zur Evaluation der Adaptivitätsmassnahmen herangezogen. Daher wird (mit Hilfe von Prokrustes-Transformationen der individuellen Karten aus verschiedenen Lernzeitpunkten) überprüft, wie stark und in welche Richtung sich die Karten nach Bearbeitung der Lernempfehlungen verändern. (Hier ist ein Kontrollgruppendesign mit Personen, die eine nicht-spezifische Lernempfehlung erhalten, angezeigt). Überproportionale Verbesserungen in den individuellen Karten der Personen der Experimentalgruppe sind dann ein starker Hinweis darauf, dass die Adaptivitätsmassnahmen greifen.²⁹ Mit den Logfiles stehen zudem Nutzungsdaten auf individueller Ebene zur Verfügung. Diese können mit den Daten der Wissensdiagnose in Beziehung gesetzt werden, wodurch auf mögliche Zusammenhänge zwischen Nutzung und Lernerfolg geschlossen werden kann.

3.4.3 Evaluation der Wirkung von PTO

PTO soll in einem ersten Testlauf unter Realbedingungen getestet werden. An diesem Punkt wird der Fokus auf die Wirkung und den Nutzen der Lernumgebung gelegt. Die gesammelten praxisbezogenen Daten helfen dann, das Angebot unter den gegebenen Bedingungen weiter zu optimieren (Tergan, 2000). Auf Ebene der Stakeholder von PTO soll anhand einer Akzeptanzanalyse festgestellt werden, wie die Lernumgebung aufgenommen wird. Hierbei geht es um die Frage, ob das System für den Benutzer in seiner subjektiven Sicht einen Mehrwert darstellt, ob er Spass hat, damit zu arbeiten, ob er es weiterempfehlen würde etc. Dazu geben die Studierenden, die mit PTO gelernt haben, über Fragebogen ein Feedback zur Lernumgebung ab. Daneben ist ebenfalls von Interesse, wie die Akzeptanz der Lernumgebung bei den Dozierenden ist, die PTO ergänzend zu ihrer Veranstaltung anbieten.

²⁹ Mit den Logfiles stehen zudem Nutzungsdaten auf individueller Ebene zur Verfügung. Diese können mit den Daten der Wissensdiagnose in Beziehung gesetzt werden, wodurch auf mögliche Zusammenhänge zwischen Nutzung und Lernerfolg geschlossen werden kann.

Weiter soll grundsätzlich der mit PTO erzielte Lernerfolg evaluiert werden. Dies wird kriteriumsorientiert geschehen, indem gemessen wird, wie die Studenten, die mit PTO lernen, bei der Abschlussprüfung abschneiden.

3.5 Schlussbemerkung

E-Learning verspricht enorme Vorteile. Der Lernprozess kann individuell dem eigenen Rhythmus und den persönlichen Vorlieben entsprechend organisiert werden. Er wird dadurch vielfältiger wie auch interessanter und gestaltet sich insgesamt effizienter. Dieser Umstand macht den Einsatz von E-Learning gerade für Universitäten interessant, wo ein hoher Grad an Selbstorganisation verlangt wird, wo aber aus Gründen der Beschränktheit der Kapazitäten bei gleichzeitig steigender Nachfrage jede Möglichkeit genutzt werden muss, die Wissensvermittlung aus dem Hörsaal in die Stuben der Studenten zu verlagern. Eine Möglichkeit den Lernprozess im E-Learning weiter zu optimieren besteht darin, Lernen adaptiv zu gestalten. Der Lerner lernt nicht mehr irgendetwas, sondern das, was er noch nicht weiss. In diesem Bericht wurde dazu eine Möglichkeit auf der Ebene von Merkmalswissen vorgestellt. Damit sich der Einsatz von E-Learning aber auch rechtfertigt, muss er bestimmten Ansprüchen genügen. Deshalb bedarf es aus der Sicht eines umfassenden Qualitätsmanagements, im gesamten Entwicklungsprozess der steten Evaluation, um die geforderten Standards erreichen und halten zu können.

3.6 Literaturverzeichnis

- Ausubel, D. P. (2000). *The acquisition and retention of knowledge*. Dodrecht: Kluwer Academic Publishers.
- Ausubel, D. P., Novak, J. D. & Hanesian, H. (1980/81). *Psychologie des Unterrichts*. 2 Bde. (2. Aufl.). Weinheim: Beltz.
- Egli, S., Schlatter, K., Streule, R. & Läge, D. (2006). A structure based expert model of the ICD-10 mental disorders. *Psychopathology* 39 (1), 1–9.
- Läge, D. & Daub, S. (1998). Die Diagnostik von Sachwissen. Zürich: *unimagazin*, 4/98, 46–50.
- Läge, D. (2001). *Ähnlichkeitsbasierte Diagnostik von Sachwissen*. Habilitationsschrift an der Philosophischen Fakultät der Universität Zürich.
- Läge, D., Oberholzer, R., Egli, S. & Streule, R. (2008). Assimilative Learning with the Aid of Cognitive Maps. *International Journal of Emerging Technologies in Learning*, 3(2), 29-33. [On-line]. Available: <http://online-journals.org/i-jet/article/view/217/203>
- Leutner, D. (1992). *Adaptive Lernsysteme. Instruktionspsychologische Grundlagen und experimentelle Analysen*. Weinheim: Psychologie Verlags Union.
- Mertens, D.M. (2005). *Research and Evaluation in Education and Psychology: Integrative Diversity with Quantitative, Qualitative, and Mixed Methods*. Thousand Oaks: Sage Publications.
- Nielsen, J. (1993). *Usability Engineering*. London: Academic Press Inc.
- Piaget, J. (1976). *Die Äquilibration der kognitiven Strukturen*. Stuttgart: Klett.
- Schweibenz, W. & Thissen, F. (2003). *Qualität im Web: Benutzerfreundliche Seiten durch Usability Evaluation*. Berlin: Springer
- Streule, R., Rüfenacht, I. & Läge D. (2003). Sachwissensdiagnostik - Was leisten Kognitive Karten? In J. Golz, F. Faul & R. Mausfeld (Hrsg.), *Experimentelle Psychologie. Abstracts der 45. Tagung experimentell arbeitender Psychologen (TeaP 2003)*. 24.-26. März, 2003, Kiel (S. 219). Lengerich: Pabst.
- Streule, R., Egli, S., Oberholzer, R. & Läge, D. (2005). Adaptive Wissensvermittlung am Beispiel der E-Learning-Umgebung „Psychopathology Taught Online“ (PTO). In D. Tavangarian & K. Nölting (Hrsg.), *Auf zu neuen Ufern! E-Learning heute und morgen* (S. 47–56). Münster / New York: Waxmann.
- Tergan, S.O. (2000). Grundlagen der Evaluation. In P. Schenkel, S.O. Tergan & A. Lottmann (Hrsg.), *Qualitätsbeurteilung multimedialer Lern- und Informationssysteme* (S. 22–51). Nürnberg: BW Bildung und Wissen.

4.

Gesamtevaluation

Abstract

Von Beginn an war Evaluation fester Bestandteil im Entwicklungsprozess der webbasierten Lernumgebung PTO (Psychopathologie Taught Online). Durch einen steten Evaluationsprozess sollte die Sicherung und Kontrolle der Qualität gewährleistet werden. Evaluation sollte helfen, dass PTO den vielfältigen Ansprüchen gerecht werden und die Kriterien für einen Einsatz in der universitären Ausbildung erfüllen konnte. Die Evaluation erfolgte systematisch während des gesamten Entwicklungsprozesses (formative und summative Evaluation) und umfasste die Ebenen Aufbau und Struktur der Lernumgebung (Produktebene), die Akzeptanz auf Seiten der Nutzer (Reaktionsebene) sowie die Effekte, welche sich mit der Arbeit mit PTO erzielen lassen (Lernebene). Ein Set an verschiedenen Methoden diente der Generierung der empirischen Daten, die beurteilt und aus denen Massnahmen zur Optimierung und Verbesserung abgeleitet wurden.

4.1 Grundlagen der Evaluation von PTO

Im Folgenden wird der gesamte Evaluationsprozess in der Entwicklung von PTO beschrieben und ausgeführt. Dies beinhaltet das Qualitätsverständnis im Rahmen des Projekts, die daraus abzuleitenden Anforderungen an die Evaluation, die eingesetzten Methoden sowie die Resultate aus den Untersuchungen. Damit soll ein umfassender Überblick über die Evaluationstätigkeiten im Rahmen des Projekts PTO gegeben werden.

4.1.1 Qualitätssicherung von PTO

Um auf universitärer Stufe von den Studierenden als Lernmittel überhaupt akzeptiert und genutzt zu werden, muss PTO hohen Qualitätsanforderungen gerecht werden. Dies setzt die Sicherstellung und Kontrolle von Qualität voraus. In welcher Form und in welchem Ausmass dies geschehen soll, ist eine grundsätzliche Frage, die sich jedem Entwicklungsteam stellt und welche je nach Projekt unterschiedlich beantwortet wird. Bevor man sich mit der Art des Ansatzes zur Qualitätssicherung beschäftigt, muss ein gemeinsam geteiltes Verständnis davon entwickelt werden, was Qualität überhaupt bedeutet. Je nach Sichtweise kann dieses unterschiedlich ausfallen. Besonders im Kontext von E-Learning ist Qualität ein Konzept, bezüglich dessen wenig Konsens herrscht und die Vorstellungen je nach Perspektive und Anspruch stark variieren. Welche Formen der Qualität für das eigene Projekt zentral sind und damit evaluiert werden sollen, ist letztlich abhängig von den Schwerpunkten, die man setzt sowie den zeitlichen, monetären und personellen Ressourcen über die man verfügt. Die Disziplin, welche sich mit der Untersuchung von Qualität beschäftigt und die nötigen Daten zur Sicherung und Kontrolle von Qualität liefert, ist die Evaluation. Diese dient der systematischen Sammlung von Daten, der Analyse und der Bewertung der Befunde mit dem Ziel der Optimierung (vgl. Böttcher, 2006).

Im Rahmen von PTO bestand von Beginn an ein ausgeprägtes Bewusstsein für die Bedeutung von Qualität. Diese wurde als multidimensionales Konstrukt wahrgenommen mit Komponenten wie technische Anforderungen, didaktische Struktur, Benutzungsfreundlichkeit, Benutzerakzeptanz und Effektivität (im Falle einer Lernumgebung entspricht diese vor allem dem Lernerfolg). Die Qualität sollte während der gesamten Entwicklung von PTO sichergestellt werden, woraus sich der Anspruch nach einem umfassenden Qualitätsmanagement ergab. Aus diesem Grund wurde innerhalb des Projektteams eine permanente Evaluationsstelle geschaffen, welche sich spezifisch mit Fragen der Qualität und Optimierung auseinandersetzte. Im Sinne eines umfassenden Qualitätsverständnisses wurde es als essentiell angesehen, dass die Lernumgebung in sämtlichen Entwicklungsphasen laufend nach verschiedenen Kriterien und mit unterschiedlichen Methoden geprüft und evaluiert wurde. Diese Analysen dienten verschiedenen Zwecken: sie boten die Grundlage für eine Standortbestimmung und Bewertung, gaben Informationen darüber,

wo die Stärken und Schwächen liegen und in welchen Bereichen sich die Lernumgebung weiter optimieren liess. Zudem sollte die Evaluation konkret zeigen, wo Sinn und Nutzen der Lernumgebung liegen. Der Hauptzweck der Evaluationsbestrebungen bestand somit darin, die Qualität von PTO auf verschiedenen Ebenen gewährleisten zu können.

4.1.2 Evaluationsziele

Im Sinne eines stimmigen Evaluationskonzepts galt es in einem ersten Schritt, die Evaluationsziele zu definieren. Für die Evaluation von PTO waren vor allem drei Aspekte von Bedeutung: Erstens das Produkt in Form der Lernumgebung. Zweitens galt es die Akzeptanz seitens der Nutzer von PTO zu erfassen. Drittens standen die Verwendung von PTO und die dabei erzielbaren Effekte (Lernen und Lernerfolg) im Fokus.

Auf Produktebene bestand ein zentrales Ziel darin, die Benutzungsfreundlichkeit sicherzustellen. Für eine webbasierte Lernumgebung ist es besonders wichtig, dass sie den Standards der Benutzungsfreundlichkeit (Usability) genügt, da viele potentielle Lerner noch wenig mit E-Learning vertraut sind und ein funktionierendes und gebrauchstaugliches System unter anderem dazu beiträgt, Berührungsängste abzubauen. Eine intuitive Bedienung ermöglicht zudem ein reibungsloses und effizientes Arbeiten.

Die erfolgreiche Einführung von PTO war weiter stark davon abhängig, ob PTO den gestellten Ansprüchen der verschiedenen Stakeholder genügt. Letztlich muss ein Lernmittel bei der Zielgruppe, für welches es konzipiert wurde, auf Akzeptanz stossen. Ist dies nicht der Fall, wird es schlicht nicht genutzt, sei das didaktische Konzept noch so ausgeklügelt und der Lernerfolg in den Augen der Entwickler noch so vielversprechend. Ein Hauptfokus der Evaluation lag schliesslich auf der Überprüfung des Adaptivitätsmechanismus sowie des Lernerfolgs. Das Adaptive Tutorielle System (ATS), welches in PTO zur Anwendung kommt, verfolgt einen innovativen Ansatz, der sich in der praktischen Anwendung noch bewähren musste.

4.1.3 Theoretische Verortung der verwendeten Evaluationsmassnahmen

Um die Evaluation systematisch und gezielt durchführen zu können, bedurfte es eines stringenten, theoriegeleiteten Konzepts. Die folgenden Ausführungen sollen in diesem Sinne die relevanten theoretischen Grundlagen für die Evaluation von PTO aufzeigen.

4.1.3.1 Evaluationsarten

Evaluationsmassnahmen differenzieren sich nach dem Zweck, dem die Massnahmen dienen und dem Zeitpunkt im Projektverlauf, an dem die Massnahmen erfolgen. Grundsätzlich lassen sich zwei Arten der Evaluation unterscheiden: formative und summative Evaluation (Tergan, 2000). Das Ziel der formativen Evaluation besteht in der

Qualitätssicherung. Sie erfolgt entwicklungsbegleitend, dient der Ermittlung von Schwachstellen und liefert damit die Datengrundlage zur Optimierung. Zum Einsatz kommen dabei in erster Linie Methoden, welche schnell relevante Daten liefern, dafür aber tendenziell pragmatisch ausgerichtet sind. Der primäre Zweck der summativen Evaluation liegt dagegen in der Kontrolle von Qualität, Wirkung und Nutzen einer Massnahme oder eines Produkts. Sie wird nach Abschluss der Implementationsphase durchgeführt und soll die Frage beantworten, ob das entwickelte Angebot den gestellten Ansprüchen genügt. Evaluationsmethoden lassen sich nicht per se einer der beiden Evaluationsarten zuordnen. Eine Befragung kann im Rahmen einer formativen Evaluation zur Qualitätssicherung, wie auch als summative Evaluation zur Qualitätskontrolle eingesetzt werden. Bei der Entwicklung von PTO wurde in der Implementationsphase zur Sicherstellung der Qualität und während der Einsatzphase zur Kontrolle der Qualität der Lernumgebung evaluiert. Die verwendeten Methoden lassen sich damit der formativen oder der summativen Evaluation zuordnen.

4.1.3.2 Produkt- und Prozessevaluation

Evaluation kann sich auf das zu entwickelnde Produkt oder auf die durchzuführende Massnahme beziehen. Hierbei spricht man von Produktevaluation. Sie kann sich weiter auf die Entwicklung und damit auf den Durchführungsprozess selber und das Verhalten der darin involvierten Personen beziehen. In diesem Fall spricht man von Prozessevaluation (Tergan, 2000). Für die Entwicklung von PTO standen Aspekte der Produktevaluation im Vordergrund. Prozessaspekte fanden ebenfalls Berücksichtigung, jedoch weit weniger ausgeprägt.

4.1.3.3 Interne und externe Evaluation

Das Projektteam setzte sich aus insgesamt fünf Personen zusammen, wobei sich je eine für die Bereiche Projektleitung, Koordination, Webdesign und Didaktik, das Adaptive Tutorielle System (ATS) sowie die Evaluation verantwortlich zeichnete (Abb. 16). Ein wichtiges Charakteristikum bei der Durchführung von Evaluationen ist die Stellung der evaluierenden Person im Entwicklungsprozess. Grundsätzlich werden interne und externe Evaluation unterschieden. Als interne Evaluation bezeichnet man die Evaluation durch eine Person, welche derjenigen Organisation angehört, die auch die Entwicklung des Produkts leitet. Ist die evaluierende Person auf Organisationsebene zugleich Teil der operativ tätigen Einheit, spricht man von Selbstevaluation (Stockmann, 2006). Im Gegensatz dazu wird bei der externen Evaluation die Evaluationstätigkeit durch eine unabhängige Organisation vorgenommen. Bei der Entwicklung von PTO war die Evaluationsstelle

zwar der operativen Einheit angegliedert, jedoch nicht in operative Tätigkeiten involviert. Demnach handelte es sich dabei um interne Evaluation.

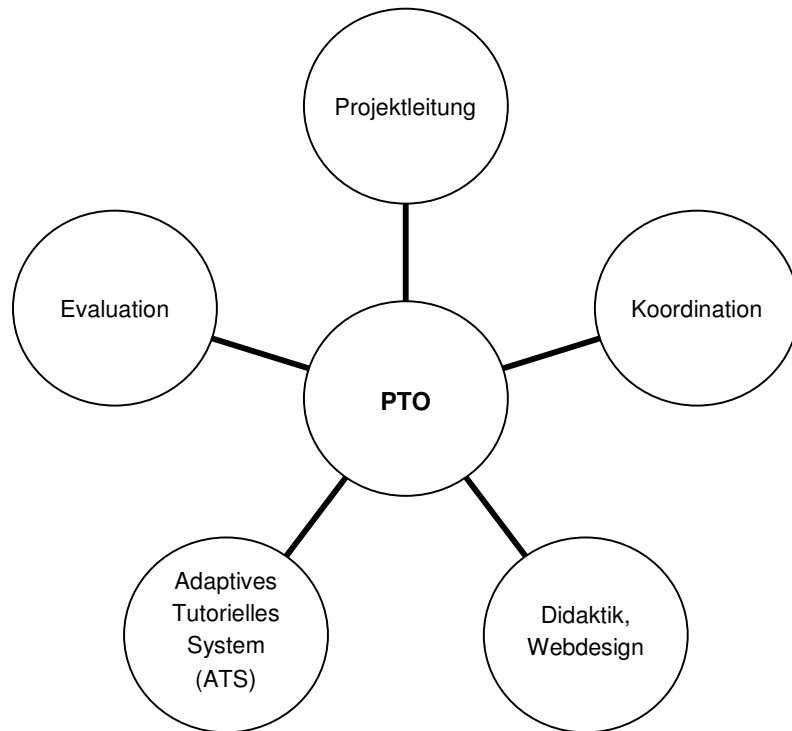


Abb. 16: Funktionen im Entwicklerteam von PTO.

Der Vorteil einer internen Evaluation liegt in der Nähe zu den Entwicklern und den übrigen Projektverantwortlichen und in der grossen Sachkenntnis der Evaluatoren. Da sich die evaluierende Person nicht direkt an der Produktion beteiligte, war genügend kritische Distanz und Unabhängigkeit zu den Inhalten und Verfahrensschritten gewahrt, um die Evaluation mit ausreichender Objektivität vornehmen zu können. Durch die Einbindung ins Team konnten die Resultate der Evaluation laufend direkt in den Entwicklungsprozess einfließen, diskutiert und bei Bedarf umgesetzt werden.

4.1.3.4 Funktionen von Evaluation

Die Funktionen der Evaluation können vielfältiger Natur sein. Rowntree (1992) unterscheidet 3 Hauptfunktionen: eine strategisch-politische Funktion, eine Kontroll- und Entscheidungsfunktion als auch eine Erkenntnisfunktion. Für PTO erscheint eine solche Sichtweise sinnvoll, legt sie doch den Fokus in erster Linie auf die für ein im universitären Umfeld eingebettetes Projekt relevanten Aspekte.

Im Sinne einer Kontroll- und Entscheidungsfunktion wird Evaluation als längerfristiger Prozess verstanden, bei dem es darum geht, laufend zu kontrollieren und Entscheidungen zu treffen, mit dem Ziel die Lernumgebung zu optimieren. Damit wird die Qualitätssicherung im Entwicklungsablauf implementiert. Für PTO bedeutete dies einen steten Optimierungsprozess mit dem Ziel, das Qualitätslevel sukzessive zu steigern.

Da PTO vom Swiss Virtual Campus (SVC) finanziell unterstützt wurde, war PTO dieser Institution gegenüber zur Offenlegung der Projektaktivitäten verpflichtet. Die Evaluationsdaten wurden vom SVC zur Überprüfung der Projektperformance genutzt und dienten als Entscheidungsgrundlage für die Sprechung neuer Gelder. Zudem sollte die Evaluationsstätigkeit den potentiellen Anwendern den Nutzen der Lernumgebung aufzeigen. Evaluation gab daher im Sinne einer strategisch-politischen Funktion Rechenschaft über den Einsatz der erhaltenen Finanzmittel und informierte die Anwender über den Nutzen der Lernumgebung.

Evaluation als Erkenntnisfunktion schliesslich betrifft die Effekte, welche durch ein Bildungsmedium, respektive ein Bildungsprogramm erzielt werden können. Im Fall von PTO betrifft dies vor allem die Akzeptanz seitens der Nutzer, als auch den Lernerfolg, den die Nutzer bei der Arbeit mit PTO zu erzielen vermögen.

4.1.3.5 Evaluationsebenen

Kirkpatrick (1998) definiert vier Ebenen der Evaluation, welche sich auf Bildungsmassnahmen anwenden lassen:

- Ebene 1: Reaktion der Teilnehmer (Akzeptanz)
- Ebene 2: Lernen (Lernerfolg)
- Ebene 3: Verhalten (Transfer)
- Ebene 4: Ergebnisse (Effekte auf die Umwelt)

Ebene 1 beinhaltet die Reaktion der Benutzer und bezieht sich damit auf den Grad der Akzeptanz gegenüber dem Bildungsangebot. Ebene 2 zielt auf den Lernerfolg ab, der beim Lernen erreicht werden kann. Ebene 3 umfasst den Transfer des theoretisch Gelernten in der praktischen Anwendung, während Ebene 4 Bezug nimmt auf die Effekte, welche sich dadurch für die Umwelt ergeben.

Da PTO mit dem Primärziel der Wissensvermittlung auf universitärer Ebene zum Einsatz kommen sollte, waren es die Ebenen 1 und 2, welche für die Evaluation besonders relevant waren. Bei PTO handelt es sich um eine Lernumgebung, womit die Lernebene und damit der Lernerfolg den Hauptanreiz bilden, weshalb sich Nutzer überhaupt mit PTO befassen. Überdies ist die Akzeptanz seitens des Benutzers eine wesentliche Voraussetzung für Lernen. Damit diese gegeben ist, braucht es ein Produkt, welches den von den

Benutzern gestellten Qualitätsansprüchen genügt. Somit standen bei der Evaluation von PTO die drei Ebenen Reaktion, Lernen und Produkt im Fokus (Abb. 17).

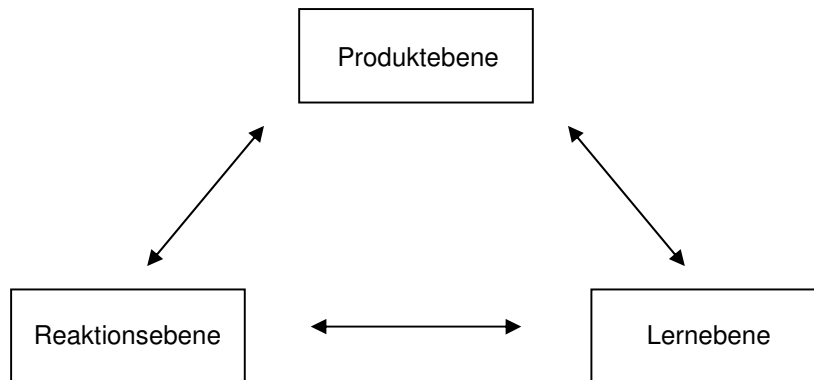


Abb. 17: Bei der Evaluation von PTO berücksichtigte Ebenen.

Auf der Produktebene standen die Merkmale des zu entwickelnden Produkts im Mittelpunkt. Produktmerkmale konnten bei der Evaluation von PTO sowohl im Prozess der Produktentwicklung als auch am fertigen Produkt evaluiert werden. Auf der Reaktionsebene wurde die Reaktion der Zielgruppe auf das Produkt untersucht. Positives Feedback war wichtig für die erfolgreiche Einführung von PTO, während negatives Feedback aufschlussreiche Anregungen für die Weiterentwicklung gab. Die Evaluation auf dieser Ebene erfolgte während der Produktentwicklung (im Rahmen eines Testeinsatzes), als auch am Ende der Entwicklungsphase. Schliesslich ging es auf der Lernebene darum, welche Effekte mit dem Produkt erreicht werden können. Für PTO waren dies vor allem der Lernprozess und der Lernerfolg. Diese wurden während der Einsatzphase untersucht.

4.2 Evaluationsmethoden und Resultate

4.2.1 Evaluationsmatrix

Anhand der oben eingeführten Kategorien Evaluationsart und Evaluationsebene, lässt sich eine Matrix aufspannen. Die Kategorie Evaluationsart umfasst die formative (Qualitätssicherung in der Entwicklungsphase) und die summativ (Qualitätskontrolle in der Einsatzphase) Evaluation. Die Kategorie Evaluationsebene beinhaltet die Produkt-, die Reaktions- sowie die Lernebene. Diese Matrix gab den Raster vor, anhand dessen PTO systematisch evaluiert werden konnte (siehe Tab. 5).

Tab. 5: Matrix zur systematischen Evaluation von PTO.

| | Produktebene | Reaktionsebene | Lernebene | Prozessevaluation: Evaluation der Arbeit im Team |
|----------------------|--|--------------------|--------------------------|---|
| Formative Evaluation | - Expert Review - Usability-Testing | - Akzeptanzanalyse | | |
| Summative Evaluation | - Überprüfung des Adaptivitätsmechanismus - Überprüfung didaktischer Elemente | - Akzeptanzanalyse | - Überprüfung Lernerfolg | |

Nachdem der konzeptuelle Raster festgelegt war, fehlten noch die entsprechenden Methoden, anhand derer evaluiert werden sollte. Die Wahl der Methoden ist grundsätzlich von den Evaluationszielen abhängig. Damit mussten für die Evaluation der Lernumgebung in einem ersten Schritt die Evaluationsziele definiert werden, aus denen sich dann die Methoden ableiten liessen.

Bezüglich der Entwicklung von PTO ergab sich gemäss dem Raster eine dreifache Zielstellung für die Evaluation: Erstens sollte die Lernumgebung als Produkt evaluiert werden. Eine solche Überprüfung umfasste die technische Funktionstüchtigkeit, Aspekte der Benutzungsfreundlichkeit etc. Weiter sollte auf Produktebene das Funktionieren des Adaptiven Tutoriellen Systems sowie der Einsatz spezifisch im Kontext von PTO entwickelter didaktischer Elemente evaluiert werden. Zweitens sollte die Akzeptanz bei den Stakeholdern (vor allem bei den Studierenden, die PTO als Lernmittel verwendeten und den Dozenten, die PTO einsetzten) erfasst werden. Drittens galt es, die Effektivität von PTO in einer Wirkungsanalyse zu prüfen. Diese umfasste den Lernerfolg, welcher bei der Arbeit mit PTO erzielt werden konnte.

Im Folgenden soll auf Grund der hier festgelegten Ziele für die einzelnen Felder des oben eingeführten Rasters ein methodischer Ansatz zur Evaluation definiert werden.

4.2.2 Formative Evaluation auf Produktebene

In der Entwicklungsphase wurde PTO mit dem Ziel evaluiert, technische Mängel sowie Mängel in der Bedienung und in der Interaktion zu eruieren. Die Interaktion zwischen

Benutzer und Lernumgebung sollte reibungslos funktionieren und der Anwender sollte diese effizient und effektiv benutzen können.

Die zur Qualitätssicherung auf Produkteebene verwendeten Methoden waren Expert Review und Usability-Testing. Expert Review ist eine expertenorientierte Methode, während Usability-Testing eine benutzerorientierte Methode darstellt. Bei Expert Review (auch heuristische Evaluation genannt) wird ein System von einem Gutachter nach Richtlinien der benutzerfreundlichen Gestaltung und software-ergonomischen Kriterien beurteilt (Schweibenz & Thissen, 2003). Resultat daraus ist eine Mängelliste, anhand derer Anpassungen mit dem Ziel vorgenommen werden können, die Benutzungsfreundlichkeit eines Systems zu optimieren.

Beim Usability-Testing (Benutzertest) testen potenzielle Nutzer ein System anhand konkreter Aufgaben. Entscheidend dabei ist, dass die Testpersonen derjenigen Personengruppe entstammen, für welche das System ursprünglich konzipiert wurde. Dies erlaubt anhand empirischer Daten einen Einblick in die tatsächliche Handhabung eines Systems und liefert Daten zur Verbesserung des Interaktionsprozesses.

Mittels Expert Review und Usability-Testing wurde die Lernumgebung im Hinblick auf ihre konkrete Nutzung durch den Benutzer untersucht. Auf Basis der erhaltenen Daten konnte die Lernumgebung iterativ im Sinne einer benutzungsfreundlichen Handhabung optimiert werden.

4.2.2.1 Evaluation der Benutzungsfreundlichkeit: Expert Review

Der Einsatz einer Expert Review empfahl sich in der Frühphase der Entwicklung von PTO, um auf diese Weise grundlegende Probleme in der Benutzung beheben zu können (ist die Darstellung einheitlich und verständlich? Gelingt die Orientierung? Werden grafische Elemente konsistent eingesetzt? Etc.). Die Lernumgebung PTO war zu diesem Zeitpunkt in einer ersten Version fertig gestellt. Eine Anzahl an Inhalten fehlte noch, jedoch müssen diese für die Durchführung einer Expert Review nicht zwingend vorhanden sein, da diese weniger auf die Inhalte als mehr auf deren Darstellung und Strukturierung fokussiert. PTO wurde in der Folge von einem Experten unter Berücksichtigung software-ergonomischer Kriterien beurteilt. Als Resultat ergab sich eine Mängelliste, die teamintern diskutiert als Grundlage zur Ableitung von Änderungsmaßnahmen diente.

Folgende Auflistung zeigt einen Auszug der Anpassungen, welche im Sinne einer benutzungsfreundlicheren Gestaltung der Benutzeroberfläche vorgenommen wurden:

- Design der Benutzeroberfläche (GUI³⁰): z.B. Anpassung Zeilenabstand; Kennzeichnung von Hyperlinks; dem Schriftkontrast förderliche farbliche Struktur des Inhaltsbereichs etc.
- Navigation: Einführung von Funktionsbuttons zum Ein- und Ausblenden von Tabelleninhalten; Einsatz von to-Top-Links etc.
- Inhalt: Einführung von Feedback-Meldungen bei kritischen Aktionen (z.B. Testabbruch); Beschriftung von Bildsteuerungselementen etc.

Mit einer Expert Review lassen sich in erster Linie „statische“ Probleme der Darstellung und Struktur eruieren. Dem geübten Blick des Experten bleiben jedoch die Probleme, die sich dem weniger erfahrenen Benutzer stellen, zum Teil verborgen. Für Problempunkte, die sich aus dem dynamischen Prozess der Interaktion zwischen „naivem“ Benutzer und Lernumgebung ergeben, eignet sich der Usability-Test als Methode.

4.2.2.2 Evaluation der Benutzungsfreundlichkeit: Usability-Testing

Auf Grund der zweiteiligen Struktur von PTO mit den Bereichen Lerninhalte und Wissensdiagnostik³¹ erschien ein separater Test der beiden Bereiche sinnvoll. Der Usability-Test I diente der Überprüfung der Benutzungstauglichkeit des Bereichs Lerninhalte. Das Testpersonenkollektiv bestand aus 4 Studierenden der Psychologie. Es wurden bewusst solche Studierende gewählt, die zwar bereits Psychologie studierten, aber noch keine Vorlesungen in Psychopathologie belegten. Dies entsprach dem zu erwartenden Profil der zukünftigen Nutzer von PTO. Das Lösen der 10 in Szenarien eingebetteten Aufgaben erforderte die konkrete Interaktion mit PTO. Auf diese Weise wurden verschiedene Aspekte von PTO wie das Design, die Navigation und die Aufbereitung und Darstellung der Inhalte geprüft. Im Anschluss an die Bearbeitung der Szenarien erfolgte ein halbstandardisiertes Interview zur Benutzungsfreundlichkeit von PTO. Zudem wurden die Probanden gebeten, den Isonorm-Fragebogen 9241/10³² auszufüllen, welcher Grundsätze zur Dialoggestaltung im Bereich der Ergonomie der Mensch-System-Interaktion definiert.

³⁰ Das Akronym GUI (Graphical User Interface) bezeichnet die grafische Schnittstelle zwischen Benutzer und System oder Maschine.

³¹ Der Bereich Lerninhalte umfasst Erläuterungen zur Phänomenologie psychischer Störungen. Der Bereich Wissensdiagnostik dagegen beinhaltet das Adaptive Tutorielle System, wo der Lerner sein Wissen testen lassen kann, Lernempfehlungen erhält und seinen Lernprozess steuert.

³² Der Isonorm-Fragebogen 9241/10 umfasst Fragen gemäß den ISO-Kriterien zur benutzerfreundlichen Dialoggestaltung.

Der Usability-Test II, welcher die Untersuchung des Bereichs Wissensdiagnostik zum Ziel hatte, umfasste insgesamt 8 Aufgaben. Erneut wurden 4 Testpersonen - nach denselben Kriterien wie im Test I - gewählt, welche die Testszenarien bearbeiteten. Im Bereich Wissensdiagnostik setzt sich der Benutzer mit Instrumenten auseinander, welche nur den wenigsten Nutzern bereits vertraut sein durften (kognitive Karten; Wissensdiagnose mittels Ähnlichkeitsurteilen etc.). Umso wichtiger erschien daher die Optimierung Usability-kritischer Aspekte, um die Nutzeranforderungen, welche durch die Komplexität und Neuartigkeit der Inhalte bereits hoch waren, nicht noch zusätzlich zu erhöhen.

Im Folgenden soll an einem konkreten Beispiel des Usability-Tests II (Überprüfung des Bereichs Wissensdiagnostik) erläutert werden, was die Evaluation gezeigt hat und welche Schlüsse daraus für die Startseite von PTO gezogen wurden. Ein Ergebnis des Tests bestand darin, dass die grundsätzliche Funktion und die Art der Verwendung des Bereichs Wissensdiagnostik von den Testpersonen zu wenig verstanden wurde. Auf Grund der Evaluation wurde in der Folge beschlossen, auf der Startseite den Titel aussagekräftiger zu wählen (siehe Abb. 18), um dadurch dem Benutzer gleich zu Beginn zu verstehen zu geben, wie die Wissensdiagnostik erfolgt (1). Im gleichen Sinne wurde der Menüpunkt „Wissensdiagnose“ in „Urteilsabgabe“ umbenannt (2). In den Tests wurde weiter festgestellt, dass die Testpersonen, ohne die Instruktionen zu lesen, die Funktionalität der Wissensdiagnostik einfach ausprobierten. Um diesem Umstand Rechnung zu tragen, wurde der Menüpunkt Wissensdiagnose, welcher ursprünglich vom Benutzer erst noch aktiviert werden musste, von Beginn an zugänglich gemacht (3). Zudem wurde auf der Startseite des Bereichs Wissensdiagnostik unter dem Link „Instruktionen“ eine Kurzanleitung angeboten, um den interessierten Benutzern die Möglichkeit zu bieten, ohne sich im Vorfeld eingehend informieren zu müssen, die Wissensdiagnostik zu nutzen (4). Schliesslich wurde der Begrüssungstext angepasst, um den Benutzer klar darüber zu informieren, was ihn im Bereich Wissensdiagnostik erwartet und wo der Nutzen in der Bearbeitung liegt (5).

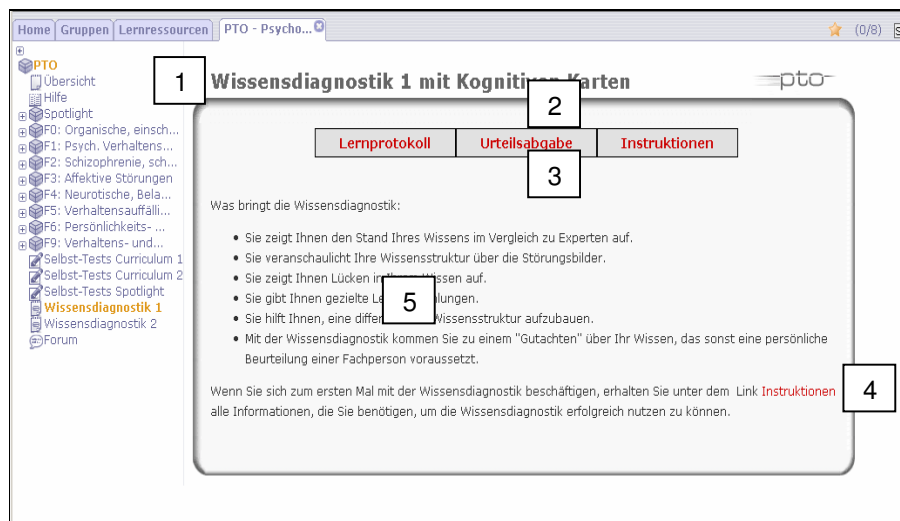


Abb. 18: Startseite des Bereichs Wissensdiagnostik vor (oben) und nach (unten) den auf Grundlage der Evaluation des Usability-Tests II vollzogenen Anpassungen.

Die beiden Usability-Tests zeigten, dass PTO im Hinblick auf die Benutzungsfreundlichkeit bereits ein Level erreicht hatte, welches einen von wenig Reibungen behafteten Umgang mit der Lernumgebung zuliess. Die Tests bestätigten damit den eingeschlagenen Weg in der Entwicklung und gaben gleichzeitig Hinweise und Ansatzpunkte für weitere Optimierung.

4.2.3 Summative Evaluation auf Produktebene

Der Bereich Wissensdiagnostik und damit das Adaptive Lernsystem von PTO ermöglichen dem Lerner die Erhebung einer Diagnose seines Wissens über die psychopathologischen Störungsbilder. Dabei werden Mängel in der Wissensstruktur detektiert, zu deren Behebung der Lerner spezifische Lernempfehlungen erhält. Es werden ihm dazu die vier am schlechtesten gewussten Störungsbilder zur Repetition empfohlen. Zudem erhält er spezifische Übungen zur Korrektur des mangelhaften Wissens. Die Evaluation des Adaptiven Tutoriellen Systems erfolgte mittels Lernexperiment. Der Adaptivitätsmechanismus wurde zuvor bereits experimentell im Labor getestet. Eine Evaluation durch ein Lernexperiment in natürlichem Setting wurde aber noch nicht durchgeführt. Ein solches Lernexperiment wurde in der Einsatzphase mit dem Ziel der Qualitätskontrolle angesetzt.

Innerhalb der Lernumgebung kommen einige speziell im Kontext von PTO entwickelte didaktische Elemente zum Einsatz. Zum einen handelt es sich dabei um eine interaktive Grafik zur Angstsymptomatik, bei der der Lerner als Lernhilfe einzelne Elemente ein- und wieder ausblenden kann. Ein weiteres solches didaktisches Element ist das Pop-up-Glossar. In PTO sind Fremd- und Fachwörter auf einfache Art und Weise abrufbar. Wird der Mauszeiger über einen solchen Begriff platziert, wird die Definition des Begriffs im Mouse-over angezeigt. Dadurch lässt sich bei Bedarf kontextspezifisch auf Informationen zurückgreifen, wodurch Unterbrechungen und Störungen im Lernprozess vermieden werden können. Als drittes Element wurde das Speed-Reading evaluiert. Diese Funktion ermöglicht es dem Lerner, mittels Knopfdruck die relevantesten Informationen markiert angezeigt zu erhalten. Dies soll helfen, den Lernprozess effizienter und effektiver zu gestalten. Diese drei didaktischen Elemente – interaktive Grafik, Pop-up-Glossar, Speed-Reading - wurden bezüglich Lernerfolg und Lerneffizienz evaluiert.

Die Evaluation des Adaptivitätsmechanismus hatte für das gesamte Projekt besondere Relevanz, da die Adaptivität im Lernprozess ein Kernelement von PTO darstellt. Die Überprüfung der didaktischen Elemente war ebenfalls von speziellem Interesse, wurden doch einige dieser Elemente spezifisch für den Einsatz in PTO entwickelt.

4.2.3.1 Evaluation des Adaptivitätsmechanismus

Die Überprüfung des Adaptivitätsmechanismus war Teil eines Lernexperiments, bei dem der Lernprozess bei der Arbeit mit PTO in einem Kontrollgruppendesign umfassend untersucht wurde. Für diesen Teil der Untersuchung lernten 28 Probanden während 7 Wochen mit PTO. Die Probanden erhielten die Aufgabe, sich „prüfungsreif“ mit den 20 Störungsbildern des Curriculums 1 von PTO auseinanderzusetzen. Sie waren bei der Wahl der Lernzeiten und der Lernörtlichkeiten frei (der Aufwand wurde mittels Lerntagebuch erfasst). Im Verlauf des Lernprozesses wurde ihr Wissen zu den in PTO behandelten psychopathologischen Störungsbildern nach Ende der ersten Lernphase (nach 6 Wochen)

sowie nach Ende der Repetitionsphase (nach 7 Wochen) erhoben. Diese Erhebungen erfolgten mit Hilfe eines Ähnlichkeitsbasierten Verfahrens, welches das strukturelle Wissen zu den Störungsbildern erfasste. Eine Eigenheit dieses Verfahrens besteht darin, dass sich anhand der erhobenen Daten eine Abbildung der Wissensstruktur in so genannten Kognitiven Karten³³ generieren lässt. In einer solchen Karte, oder über den Vergleich mehrerer solcher Karten einer Person zu verschiedenen Zeitpunkten, lässt sich über die Änderung der Struktur, respektive die Veränderung der Positionen der einzelnen gelernten Störungsbilder, der Lernprozess visualisieren. In der Positionsänderung der 4 repetitierten Störungsbilder in den Karten zeigte sich die Wirkung des Adaptivitätsmechanismus. Die Wirkung konnte dabei zum einen auf individueller Ebene anhand der Positionsänderungen in der Karte aber auch numerisch auf Stichprobenebene über die Änderungen im ObjectLoss (dem Mass der Abweichung von der Optimalposition) untersucht werden. Das Vorgehen und die Resultate zur Evaluation des Adaptivitätsmechanismus sind im Kapitel 5 ausführlich beschrieben. Nachfolgende Abb. 19 zeigt auf individueller Ebene anhand einer Kognitiven Karte beispielhaft den Lernprozess der Testperson E5. Diese erhielt nach der ersten Lernphase 4 Störungsbilder (in der Abbildung mit der Laufnummer 1, 9, 17 und 19 versehen) zur Repetition empfohlen. In der hier präsentierten Karte wurden die repetitierten Störungsbilder in die auf Expertenwissen basierende Struktur der 20 Störungsbilder eingefügt. Dadurch wird ersichtlich, in welchem Ausmass sich diese durch die Repetition den Expertenpositionen angenähert haben. Es zeigt sich deutlich, dass sich die repetitierten Störungsbilder vom Zeitpunkt t2 (vor der Repetition) nach t3 (nach der Repetition) Richtung Expertenposition bewegen (Abb. 19 verdeutlicht diese Verschiebung).

³³ Eine Kognitive Karte zeigt eine strukturelle Ansicht des Wissens über ein bestimmtes Inhaltsgebiet. Als Basis für das Erstellen der Karte dienen Ähnlichkeitsurteile, welche die jeweilige Person zwischen je zwei in der Karte abzubildenden Störungsbildern abgibt.

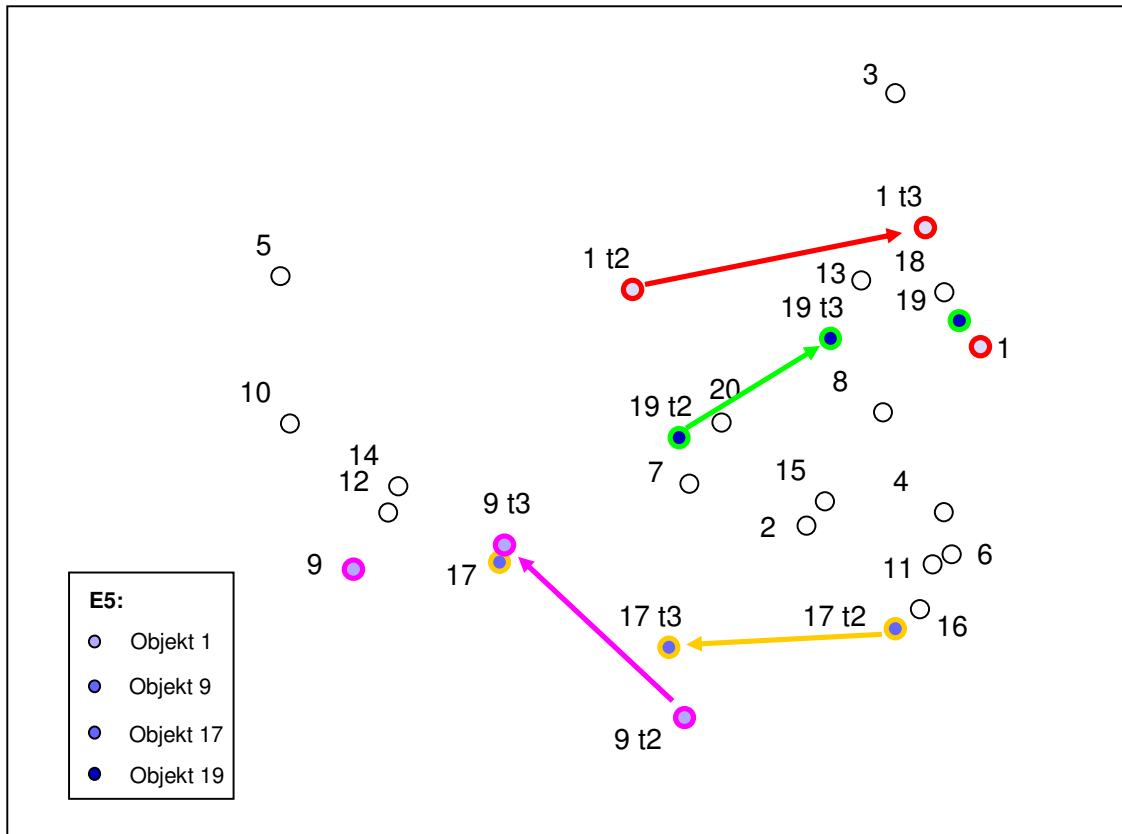


Abb. 19: Positionsänderungen von t2 (vor der Repetition) nach t3 (nach der Repetition) der 4 repetierten Störungsbilder der Testperson E5 in Bezug zu den Expertenpositionen.

Bezüglich des Ausmasses, mit dem sich die 4 repetierten Objekte bewegen, bestehen nur geringfügige Unterschiede. Die Verbesserungen fallen für alle repetierten Störungsbilder ähnlich stark aus. Zum Zeitpunkt t3 positionieren sie sich in deutlicher Nähe zur Expertenposition.

4.2.3.2 Evaluation didaktischer Elemente

Die gesamten Resultate zur Evaluation der didaktischen Elemente finden sich ausführlich beschrieben im Kapitel 6. Im Folgenden soll die Evaluation am Beispiel des didaktischen Elements interaktive Grafik erläutert werden. Diese Grafik enthält die 14 massgeblichen Angst- und Paniksymptome in 4 Oberkategorien unterteilt. Untersucht werden sollte das Lernen mit der Angstkreis-Grafik im Vergleich zum herkömmlichen, buchbasierten Lernen. Das Lernziel im Versuch bestand im Aufzählen der Symptomklassen und der zugehörigen Symptome sowie in der korrekten Zuteilung der Symptome zu den Kategorien. Getestet wurden zwei Lerngruppen, wobei je eine mit einem der beiden Formate lernte. Eine Zeitvorgabe bestand nicht, die Probanden wurden instruiert, die Inhalte solange zu

lernen, bis sie subjektiv der Überzeugung waren, diese vollständig wiedergeben zu können.

Untersucht wurde bezüglich Lerneffizienz (gemessen über die Lernzeit) und Lernerfolg (gemessen über die Korrektheit der Resultate). Insgesamt wurden 24 Personen getestet. Experimental- und Kontrollgruppe enthielten je 12 Personen. Als Probanden wurden Personen aus der potenziellen Nutzergruppe von PTO gewählt.

Abb. 20 zeigt die Lerndauer für die Experimentalgruppe, welche anhand der interaktiven Grafik gelernt hatte, sowie der Kontrollgruppe, welche sich die Lerninhalte mittels Lehrbuch angeeignet hatte. Die Experimentalgruppe benötigte durchschnittlich 361 Sekunden (Stdv: 109 s), die Kontrollgruppe dagegen 658 Sekunden (Stdv: 377 s), für das Lernen der Inhalte. Der Unterschied wurde auf dem 5%-Niveau signifikant ($t(13) = -2.62, p < .05$).

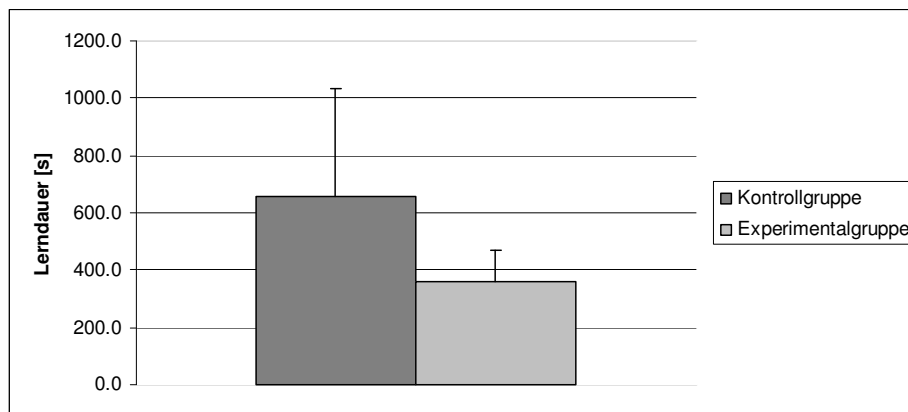


Abb. 20: Durchschnittliche Lerndauer der Kontrollgruppe (gelernt mit Lehrbuch) sowie der Experimentalgruppe (gelernt mit interaktiver Grafik).

Abb. 21 zeigt die durchschnittliche Anzahl korrekter Antworten für die Experimental- und die Kontrollgruppe. Die Gesamtzahl möglicher korrekter Antworten lag bei 18. Die Experimentalgruppe erzielte durchschnittlich 16.3 korrekte Antworten (Stdv: 2), während deren Zahl bei der Kontrollgruppe 15.8 (Stdv: 1.6) betrug. Der gefundene Unterschied zwischen den beiden Gruppen wurde jedoch nicht signifikant ($t(22) = 0.56, p > .05$).

Bezüglich der Anzahl nicht korrekter Antworten (falsche und nicht gewusste Antworten) konnte zwischen den Gruppen kein signifikanter Unterschied festgestellt werden.

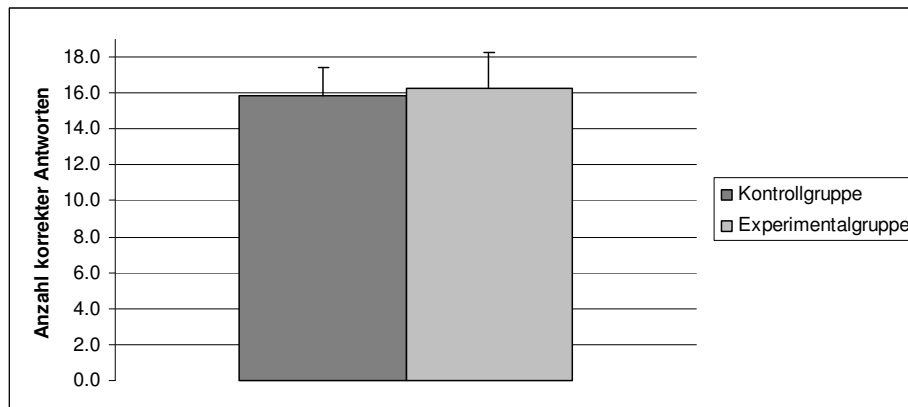


Abb. 21: Durchschnittliche Anzahl korrekter Antworten der Kontrollgruppe (gelernt mit Lehrbuch) sowie der Experimentalgruppe (gelernt mit interaktiver Grafik).

Die Resultate zeigen, dass diejenigen Personen, welche mit der interaktiven Grafik lernten, bei durchschnittlich gleicher Anzahl korrekter Antworten signifikant weniger Zeit benötigten als die Lehrbuchlerner. Für die beiden weiteren untersuchten didaktischen Elemente, Pop-up-Glossar und Speed Reading, wurden im Vergleich zum Lehrbuchlernen ebenfalls signifikant bessere Ergebnisse bezüglich des Lernerfolgs erzielt.

4.2.4 Formative und summative Evaluation auf Reaktionsebene

Ein wichtiges Evaluationsziel bestand in der Erfassung der Benutzerakzeptanz. Diese war für den erfolgreichen Einsatz von PTO zentral. Die Akzeptanz einer Lernumgebung ist von verschiedenen Faktoren wie der Lernmotivation, der grundsätzlichen Einstellung gegenüber Computern und E-Learning, dem erwarteten Nutzen, der Benutzungsfreundlichkeit etc. abhängig. Bei einer Befragung müssen daher verschiedene Faktoren berücksichtigt werden. Die Akzeptanzerhebung während der Entwicklungsphase (formative Evaluation) sollte zeigen, in welchen Bereichen PTO in den Augen der Benutzer gut abschneidet und gleichzeitig wertvolle Hinweise auf Optimierungspotenzial zeigen. Eine Erhebung während der Einsatzphase sollte im Sinne einer Qualitätskontrolle auf Reaktionsebene die Benutzerakzeptanz des fertigen Produkts erfassen. Der Benutzerakzeptanz wurde insgesamt grosse Bedeutung beigemessen. Der erfolgreiche Einsatz einer webbasierten Lernumgebung steht und fällt letztlich mit der Zufriedenheit bei der Zielgruppe.

4.2.4.1 Evaluation der Benutzerakzeptanz

Akzeptanzerhebungen wurden zu verschiedenen Zeitpunkten im Entwicklungsprozess durchgeführt. Eine erste Erhebung erfolgte als PTO testhalber während eines Semesters

den Studierenden der Psychopathologie der Universität Zürich zur Verfügung gestellt wurde.

Während dieser Zeit wurde zu zwei unterschiedlichen Zeitpunkten die Akzeptanz von PTO erhoben. Die erste Erhebung richtete sich an die Studierenden und erfolgte über einen online-Fragebogen, nachdem diese bereits während 4 Wochen auf PTO zugreifen konnten. Diese Erhebung war spezifisch auf den Bereich Lerninhalte ausgerichtet. Gefragt wurde nach dem Nutzungsverhalten der Studierenden, dem Nutzen der verschiedenen Strukturelemente und Materialien von PTO, der Benutzerfreundlichkeit der Lernumgebung, der generellen Einschätzung von PTO sowie nach konkreten Verbesserungsvorschlägen. Die Befragung sollte beispielsweise Informationen darüber liefern, ob und wie PTO von den Studierenden genutzt wird, wie Aufbau, Struktur und Gestaltung eingeschätzt wurden und ob PTO für die Studierenden einen effektiven Mehrwert darstellte.

Die zweite Erhebung erfolgte weitere 4 Wochen später wiederum per online-Fragebogen. Bei dieser Befragung stand der Bereich Wissensdiagnostik im Fokus. Gefragt wurden die Studierenden bezüglich der Nutzung, der Verständlichkeit und der Benutzerfreundlichkeit des Bereichs Wissensdiagnostik. Zudem wurde analog zur ersten Akzeptanzbefragung die generelle Einschätzung von PTO erfragt, um dadurch im Vergleich mit den Resultaten der ersten Erhebung eine Aussage über PTO im zeitlichen Verlauf machen zu können. Spezifisch für den Bereich Wissensdiagnostik sollte die Erhebung Informationen darüber liefern, ob die Studierenden Aufbau, Struktur und Darstellung des Bereichs verstanden, das Verfahren zur Wissensdiagnose sowie das Feedback zur Diagnose akzeptierten und ob die Wissensdiagnostik generell einen Mehrwert für sie darstellte.

Da der Gebrauch von PTO auf freiwilliger Basis erfolgte und keine äusseren Anreize, wie beispielsweise eine Prüfung am Ende des Kurses, bestanden, war die Rücklaufquote erwartungsgemäss relativ gering. Insgesamt beteiligten sich 13 Personen (9 valide Antwortbogen) an der Befragung zum Bereich Lerninhalte sowie 9 Personen (5 valide Antwortbogen) an der Befragung zum Bereich Wissensdiagnostik.

Die dritte Erhebung richtete sich an die beiden Dozenten, in deren Vorlesung PTO im Rahmen eines Blenden Learning-Angebots eingesetzt wurde. In einem offenen Interview wurden diese zu PTO und dem Einsatz der Lernumgebung im Rahmen ihrer Vorlesung befragt. Dabei bestand das Ziel festzustellen, wie PTO von Seiten der Dozenten in die Vorlesung eingebunden wurde, wo sie das Potential von PTO sahen, wie sie sich dessen optimalen Einsatz in der Vorlesung vorstellten und wo sie Risiken beziehungsweise Verbesserungsmöglichkeiten ausmachten.

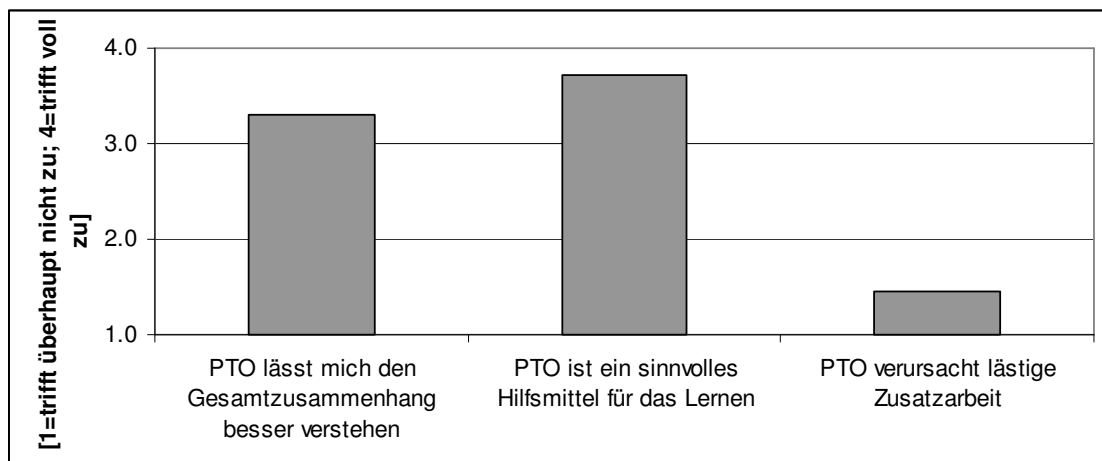
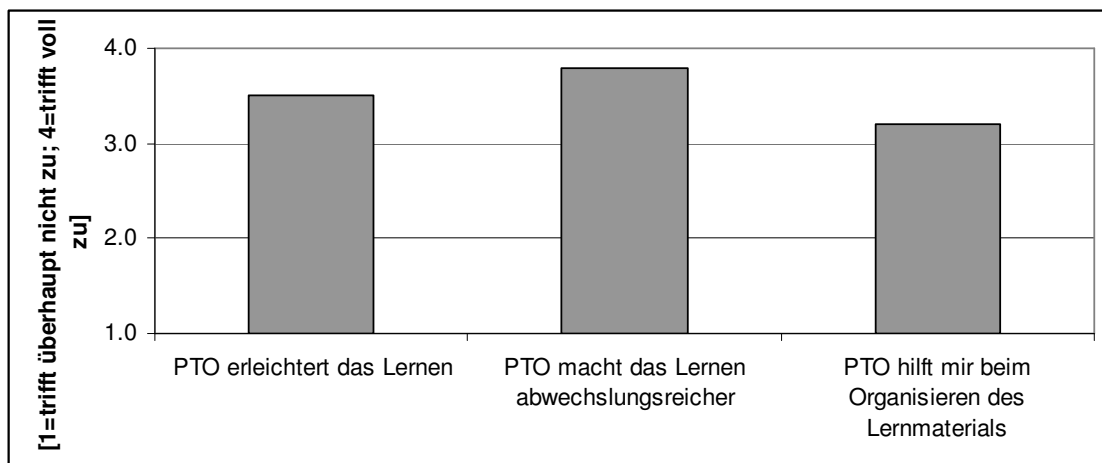
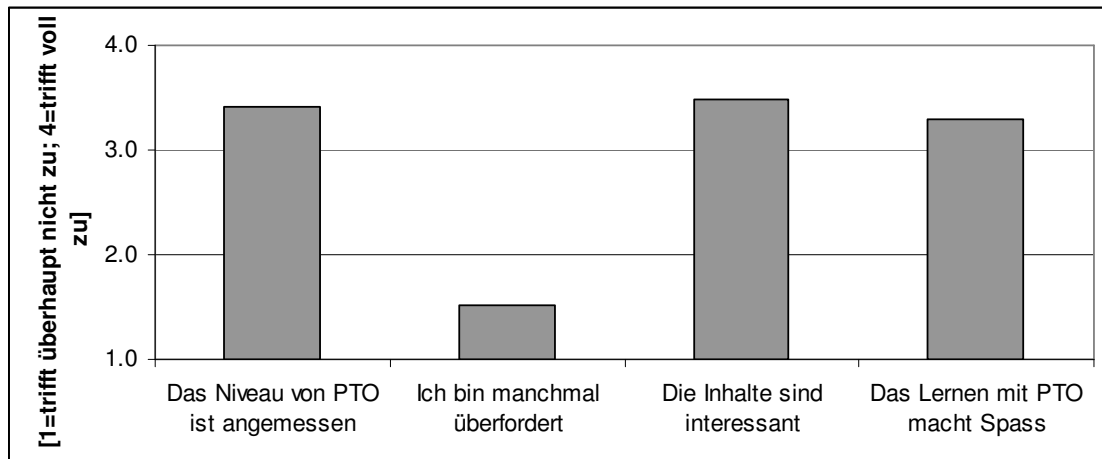
Eine vierte Akzeptanzerhebung fand während der Durchführung des im Kapitel 5 detailliert erläuterten Lernexperiments statt. Bei den Teilnehmern des Lernexperiments handelte es sich um diejenigen Personen, welche sich bis zu diesem Zeitpunkt am intensivsten mit PTO auseinandergesetzt hatten. Umso aufschlussreicher waren damit die Ergebnisse

der Befragung. Art und Struktur der Erhebung erfolgten analog zu den ersten beiden Akzeptanzbefragungen, um dadurch einen Vergleich der Resultate gewährleisten zu können. Gesamthaft nahmen 9 Personen der Experimentalgruppe an der Befragung zu den Bereichen Lerninhalte und Wissensdiagnostik teil. 10 Personen der Kontrollgruppe füllten den Fragebogen zur Wissensdiagnostik aus.

Im Rahmen einer Blended-Learning-Veranstaltung (Präsenzvorlesung zur Psychopathologie kombiniert mit der Arbeit mit PTO) wurden 3 Übungen eingesetzt, bei denen die Studierenden in Kleingruppen verschiedene Aufgaben zu bearbeiten hatten. Das Bearbeiten dieser Übungen erforderte unter anderem die Auseinandersetzung mit PTO. Dieses Konzept der Kombination von Präsenzvorlesung, selbständiger Arbeit mit PTO sowie Bearbeitung der Übungen wurde in der Folge mittels online-Fragebogen bezüglich Akzeptanz seitens der Teilnehmer erhoben. An dieser Befragung beteiligten sich 69 Personen (50 valide Fragebogen). Die hohe Zahl an Teilnehmenden ist darauf zurückzuführen, dass das Lösen der 3 Übungen zur Erfüllung der Anforderungen des entsprechenden Kurses zwingend war.

Weiter wurde PTO den Psychologiestudierenden der Partneruniversität Salzburg zur Verfügung gestellt. Diese wurden ebenfalls mittels online-Fragebogen zu PTO befragt. Im Rahmen des Einsatzes von PTO an der Partneruniversität Fribourg fand eine weitere Akzeptanzbefragung statt. Diese erfolgte in verkürzter Form und beinhaltete in offener Form Fragen dazu, wo die Studierenden positive Aspekte von PTO ausmachten und welche Verbesserungsmöglichkeiten sie sahen. In Salzburg und Fribourg nahmen je 9 Personen an der Befragung teil.

Die folgenden Grafiken zeigen Antworten zu Fragen zum Bereich Lerninhalte, wie sie im online-Fragebogen der Akzeptanzerhebung gestellt wurden. Die Beantwortung der Fragen erfolgte anhand von vierstufigen Ratingskalen (1=trifft überhaupt nicht zu; 2=trifft eher nicht zu; 3=trifft eher zu; 4=trifft voll zu). Die Antworten stammen von insgesamt 30 Studierenden der Universitäten Zürich und Salzburg, die mit PTO gearbeitet hatten.



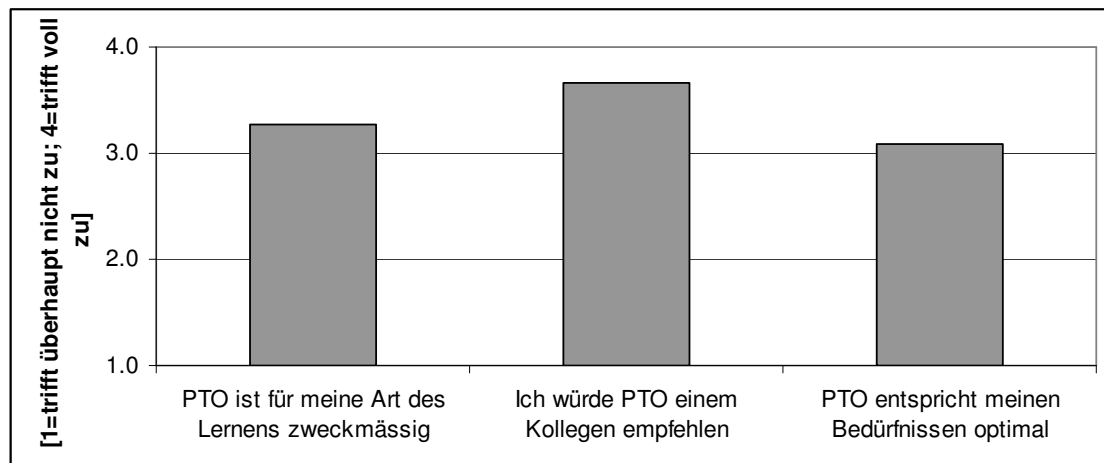


Abb. 22: Resultate der Akzeptanzanalyse zum Bereich Lerninhalte (n=30).

Die Befragungen legten dar, dass PTO seitens der Benutzer grundsätzlich auf grosse Akzeptanz stösst. Insbesondere die Auswahl und die Art der Präsentation der Inhalte wurden sehr positiv bewertet, was zeigte, dass PTO in der Wissensvermittlung aus Sicht der Benutzer einen echten Mehrwert bietet. Die Studierenden äusserten sich sehr positiv zu den verwendeten Materialien sowie der Art und Weise wie diese kombiniert eingesetzt werden. Die Benutzungsfreundlichkeit wurde allgemein als sehr gut eingestuft. Die spezifischen Fragen zur Navigation zeigten, dass diese sehr gut bewertet wurden und die Struktur und das Design von PTO es den Studenten ermöglicht, sich ohne Probleme zu orientieren und mit der Lernumgebung zu interagieren.

Auch bezüglich der generellen Einschätzung schnitt PTO sehr gut ab. So wurde das Niveau als angemessen eingestuft. Die Inhalte entsprachen – zumindest aus Sicht der Studierenden – dem universitären Niveau. Die Studierenden fühlten sich nicht überfordert. Die Inhalte wurden generell als interessant bewertet und das Lernen mit PTO bereitet Spass. Weiter gaben sie an, dass PTO das Lernen erleichtert, es abwechslungsreicher gestaltet und sie dabei unterstützt, den Gesamtzusammenhang besser zu verstehen. Die Beschäftigung mit PTO wurde nicht als mühsame Zusatzarbeit wahrgenommen und entsprach den Bedürfnissen der Studierenden stark. PTO schien damit die Möglichkeiten, die eine multimedial aufbereitete Lernumgebung mit sich bringt, im Sinne der Bedürfnisse der Studierenden zu nutzen.

Die Akzeptanzbefragung zeigte weiter, dass die Studierenden mit dem Bereich Wissensdiagnostik teilweise Schwierigkeiten hatten. Die Möglichkeit einer Wissensdiagnose und die Idee des adaptiven Lernens wurden sehr positiv bewertet. Die Art der Diagnose mittels Ähnlichkeitsurteilen und die Interpretation der resultierenden Karte waren vorher kaum einem der Nutzer vertraut und führten teilweise zu Verwirrung. Weiter wurde die Länge der Urteilsabgabe kritisiert. Daher wurden in der Folge die Instruktionstexte ange-

passt und strukturelle Änderungen vorgenommen, um das Verständnis der Wissensdiagnostik klarer und deren Gebrauch einfacher zu gestalten.

In den Interviews mit den Dozenten konnte der Blended-Learning-Einsatz von PTO aus Sicht der Vorlesungsveranstalter erfasst werden. Die Dozenten sahen einen gewichtigen Vorteil von PTO darin, dass die Studierenden den Stoff in der Vorlesung vermittelt bekommen, diesen aber zu Hause selbständig vor- und nachbereiten können. Insbesondere die multimediale Aufbereitung helfe den Studierenden, das Gelernte besser zu verankern und führe zu einem umfassenderen Verständnis. Zudem finde eine Entlastung der Vorlesung statt, da der Dozent die Vermittlung bestimmter Inhalte an PTO delegieren könne. Die Studierenden verfügten bereits über ein Basiswissen, wenn sie in die Vorlesung kämen, wodurch dort auf spezifische Aspekte eingegangen werden könne, die Vorlesung stärkeren Frage-Antwort-Charakter erhalte und die Interaktion zwischen Dozent und Besucher allgemein grösser würde. Überdies könnten in der Vorlesung Anwendungsbeispiele und Erfahrungsberichte präsentiert werden, wodurch sich für die Studierenden ein konkreter Bezug zur Materie herstellen liesse.

Die Befragung zu den Blended-Learning-Übungen zeigte, wo die Studierenden die Vor- und Nachteile in Bezug auf die 3 zu bearbeitenden Übungen sahen. Weiter wurde nach der Akzeptanz in Bezug auf organisatorische Aspekte wie Gruppengrösse, Zeitaufwand etc. gefragt. Die Befragung ergab damit wichtige Information darüber, welche Aspekte der Gestaltung der Übungen gut funktionierten und wo im Sinne einer höheren Nutzerakzeptanz Verbesserungspotenzial vorhanden war.

4.2.5 Summative Evaluation auf Lernebene

Ein weiteres Ziel bei der Evaluation bestand darin, die Effektivität von PTO in einer Wirkungsanalyse zu prüfen. Dazu sollte der Lernerfolg erfasst werden, welcher bei der Arbeit mit PTO erzielt werden kann. Der erreichbare Lernerfolg stellt gemeinhin das für die Benutzer entscheidende Qualitätskriterium einer Lernumgebung dar. Umso wichtiger ist die umfassende Evaluation des Lernerfolgs, kann diese doch Daten liefern, welche potenziellen Nutzern gegenüber als Argument zur Nutzung der Lernumgebung vorgebracht werden können. Die Absicht des Lernexperiments bestand darin, das Wissen über die psychopathologischen Störungsbilder im Verlauf des Lernprozesses auf Individual- wie auch auf Gruppenebene zu überprüfen. Bei der Evaluation mittels Lernexperiment handelt es sich um Qualitätskontrolle auf Lernebene.

In der Entwicklungsphase wurde auf der Lernebene auf eine formative Evaluation verzichtet. Die Überprüfung des Lernerfolgs setzt eine vollfunktionstüchtige und alle Lerninhalte umfassende Lernumgebung voraus. Naturgemäss waren die Voraussetzungen in der Entwicklungsphase hierfür nicht gegeben.

4.2.5.1 Evaluation des Lernerfolgs

Die Überprüfung des Lernerfolgs war Teil eines Lernexperiments, welches den Lernprozess im Rahmen von PTO umfassend untersuchte (siehe Kapitel 5). Für diesen Teil des Experiments lernten 28 Personen während 7 Wochen mit PTO. Während der ersten 6 Wochen bearbeiteten sie die 20 Störungsbilder des Curriculums 1 von PTO. Danach erfolgte eine Wissensdiagnose. Die 4 dabei am schlechtesten gewussten Störungsbilder wurden in der Folge in der 7. Lernwoche repetiert. Vor Lernbeginn, nach 6 Wochen und nach der 7. Woche, am Ende der Repetitionsphase, wurde eine Wissensdiagnose durchgeführt. Ziel war es, die Lernfortschritte auf der Ebene des deklarativen, als auch des strukturellen Wissens zu erfassen. Die Erfassung des deklarativen Wissens erfolgte über einen Multiple-Choice-Test, wohingegen das strukturelle Wissen über ein ähnlichkeitsbasiertes Verfahren erhoben wurde. Auf Basis der Ähnlichkeitsdaten liess sich das strukturelle Wissen in Kognitiven Karten abbilden und veranschaulichen. Da das Experiment im Kontrollgruppendesign angelegt war, wurden die Multiple-Choice-Tests sowie die Erhebungen mittels Ähnlichkeitsurteilen von der Kontrollgruppe ebenfalls erhoben. Mit Hilfe des Lernexperiments konnte der Lernerfolg auf deklarativer Ebene wie auch auf Strukturebene nachgewiesen werden. Dies zum einen numerisch auf Grund signifikanter Verbesserung von Kennwerten, zum andern aber auch grafisch durch Visualisierung des Lernprozesses in den Kognitiven Karten.

Abb. 23 veranschaulicht beispielhaft den Lernprozess anhand einer LOMM-Karte (Loss Oriented Meta Map). In einer solchen Karte sind die einzelnen Testpersonen durch jeweils einen Punkt repräsentiert. Jene Personen, die eine ähnliche Wissensstruktur aufweisen, werden in der Karte auch nahe beieinander positioniert. Die Karte zeigt die Positionen der Lernergruppe vor Lernbeginn (t1), nach der Bearbeitung der 20 Störungsbilder (t2) sowie nach der Repetitionsphase (t3). Die Expertenstruktur, welche die Ziel- oder Optimalstruktur darstellt, wird durch einen weiteren Punkt repräsentiert. Die Cluster kennzeichnen die verschiedenen Erhebungszeitpunkte.

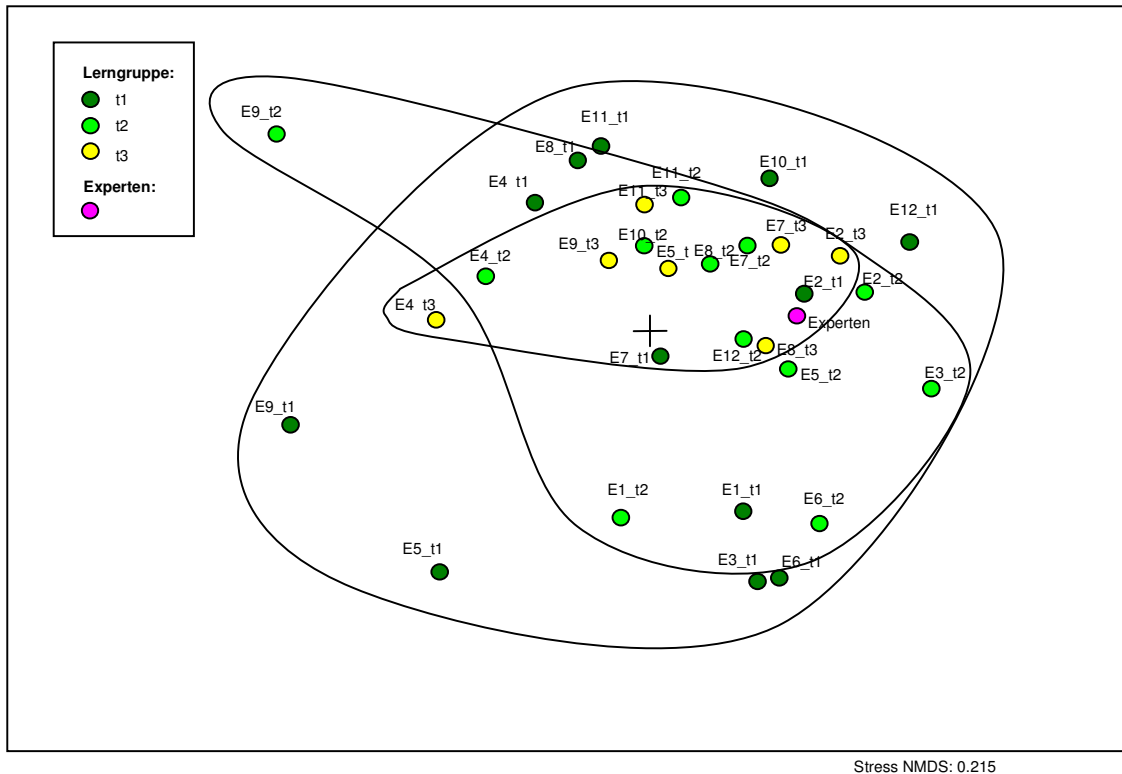


Abb. 23: LOMM-Karte mit den Positionen der Lernergruppen. Die Cluster kennzeichnen die drei Erhebungszeitpunkte t1, t2 und t3.

Es zeigt sich die starke Tendenz, dass sich die Probanden durch die Lernphase der Expertenposition annähern und sich damit in Richtung Expertenstruktur verbessern. Der Cluster zum Zeitpunkt t2 weist mit der Person E9 einen markanten Ausreisser auf. Gesamthaft betrachtet nähern sich die Wissensstrukturen der Lernergruppe durch das Lernen mit PTO aber sukzessive derjenigen der Experten an.

4.2.6 Prozessorientierte Methoden

Zusätzlich zu den am Produkt Lernumgebung ausgerichteten Evaluationsmassnahmen wurde der Entwicklungsprozess, respektive die darin involvierten Personen und ihre Arbeit im Projekt evaluiert. Das Entwicklerteam wurde hierzu mit dem Fragebogen zur Arbeit im Team (FAT) (Kauffeld, 2004) befragt. Dieses Diagnoseinstrument dient der Ableitung des Teamentwicklungsbedarfs. Befragt wurden die Projektmitglieder bezüglich der Faktoren Zielorientierung, Aufgabenbewältigung, Zusammenhalt sowie Verantwortungsübernahme. Auf die Befragung der Projektpartner oder weiterer an der Entwicklung Beteiligter wurde verzichtet, da sich der FAT an eng zusammenarbeitende Teams richtet. Da dieser Grad an Zusammenarbeit nur innerhalb des Kernteams erreicht wurde, schien eine Befragung der Partner wenig sinnvoll.

Abb. 24 zeigt die eingesetzten Evaluationsmethoden in der chronologischen Abfolge ihrer Anwendung. Dadurch erschliesst sich das logisch-zweckmässige Zusammenspiel der einzelnen Methoden im Entwicklungsprozess.

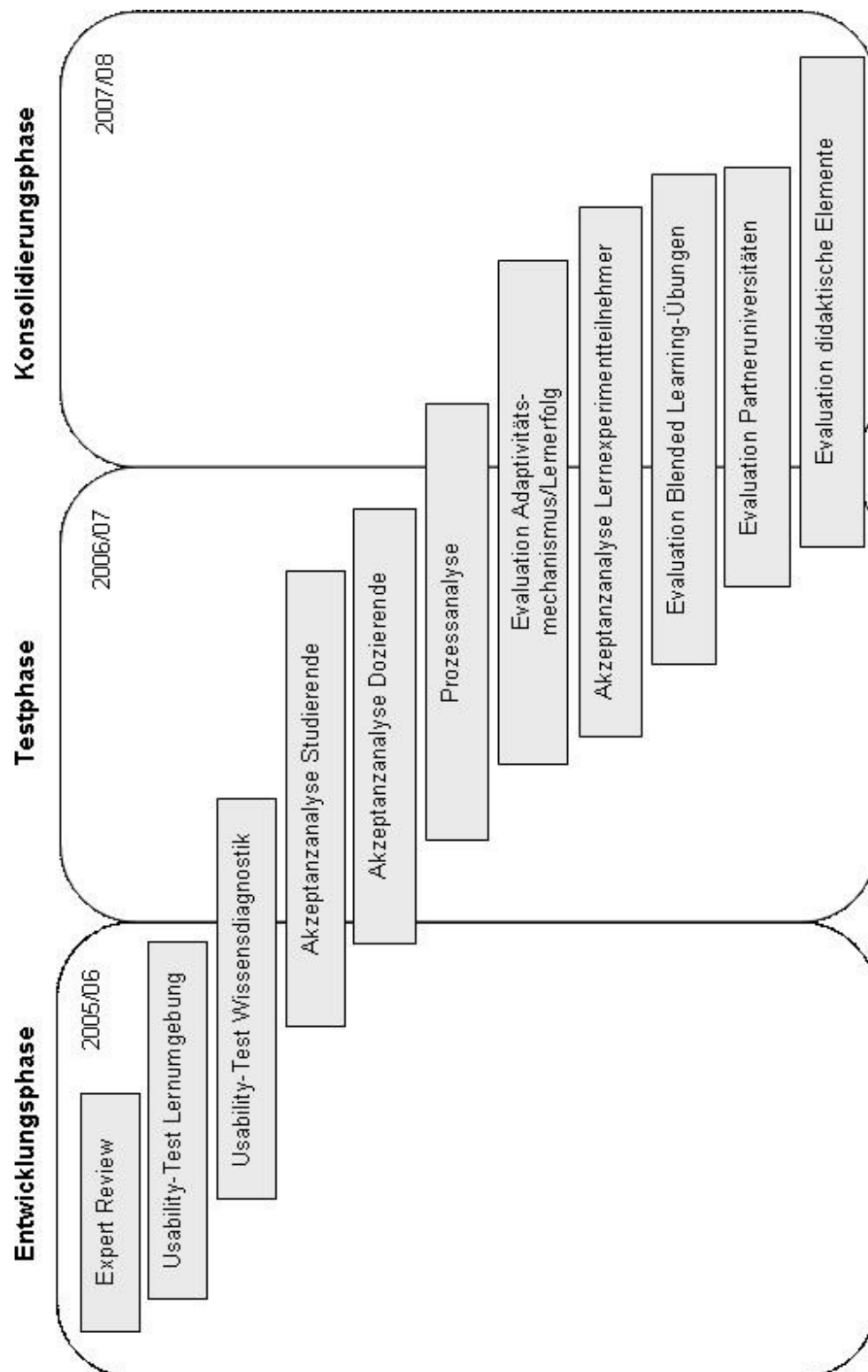


Abb. 24: Evaluationsmassnahmen im zeitlichen Verlauf.

4.3 Diskussion

Evaluation war von Beginn an fester Bestandteil im Rahmen der Entwicklung von PTO. Für die Durchführung der verschiedenen Evaluationsaktivitäten, wurde hierzu eine permanente Stelle innerhalb des Entwicklerteams geschaffen. Evaluiert wurde mit dem Ziel der Qualitätssicherung (formative Evaluation) als auch der Qualitätskontrolle (summative Evaluation). Die Massnahmen bezogen sich dabei auf die Struktur und den Aufbau der Lernumgebung selber (Produktebene), auf die Akzeptanz von Seiten der Benutzer (Reaktionsebene) sowie auf die Effekte, die sich durch die Arbeit mit PTO erzielen liessen (Lernebene). Mit den Kategorien Art und Ebene der Evaluation war ein Raster vorgegeben, anhand dessen die Lernumgebung systematisch getestet und evaluiert werden konnte. Die vorgenommenen Massnahmen hierzu wurden in diesem Bericht vorgestellt und ausgewählte Resultate präsentiert. Wo die Resultate an anderer Stelle in umfassenderer Form ausgeführt werden, wurde an entsprechender Stelle darauf verwiesen.

Die Evaluationsmassnahmen dienten der Gewinnung von Informationen, welche in der Folge einer Bewertung unterzogen wurden und damit die empirische Datengrundlage für Entscheidungen zur Optimierung von PTO lieferten. Dank Expert Review und Usability-Tests liess sich die Benutzeroberfläche optimieren und die Interaktion zwischen Benutzer und PTO verbessern. Die Akzeptanzanalysen zeigten auf verschiedenen Ebenen, wie die Nutzer die Arbeit mit PTO erlebten, welche Aspekte sie positiv beurteilten und wo sie noch Optimierungspotenzial sahen. Durch diese Befragungen, welche regelmässig durchgeführt wurden, konnte PTO im Sinne der Anwender angepasst werden. Als gewinnbringend hat sich zudem erwiesen, dass neben der Anwendersicht der Studierenden auch die Perspektive der Dozenten, welche PTO in ihren Vorlesungen im Rahmen von Blended-Learning einsetzten, in Interviews erfasst wurde.

Eine für das Projekt besonders relevante Evaluation war diejenige des Adaptivitätsmechanismus sowie des Lernerfolgs, welche anhand eines Lernexperiments vorgenommen wurden. Mit diesem Experiment konnte der Lernerfolg, der mit der Arbeit mit PTO erzielt werden kann, auf Ebene des Fakten- als auch auf Ebene des strukturellen Wissens nachgewiesen werden. Zudem zeigte sich das Funktionieren des Adaptivitätsmechanismus auf individueller Ebene.

Eine weitere Evaluation zielte auf drei didaktische Elemente, welche spezifisch für den Einsatz in PTO entwickelt wurden. Diese wurden in einer Vergleichsstudie in Bezug auf Lernerfolg mit dem klassischen Lehrbuchlernen verglichen. Der Lernerfolg erwies sich beim Lernen auf Grundlage der didaktischen Elemente als höher.

Probleme bei der Evaluation gab es dort, wo die Ergebnisse von der freiwilligen Teilnahme der Testpersonen abhängig waren. Dies zeigte sich beispielsweise beim testweisen Einsatz von PTO. Die Bearbeitung von PTO hatte für die Studierenden keine unmittelbare Prüfungsrelevanz und damit wahrscheinlich geringere Priorität. Dies war der mutmass-

liche Grund, weshalb die Rücklaufquote in der online-Befragung eher niedrig ausfiel. Im Lernexperiment zeigte sich ein ähnliches Problem, welches jedoch sämtliche auf einen längeren Zeitraum angelegte Experimente betrifft: der Ausfall von Probanden im Testverlauf. Die Probanden mussten sich über sieben Wochen hinweg intensiv mit PTO auseinandersetzen. Dies setzte ein hohes Mass an Selbstorganisation voraus. Die Teilnahme am Experiment war zudem freiwillig, zeitintensiv und ein vorzeitiger Ausstieg war mit keinerlei negativen Konsequenzen verbunden. Dementsprechend waren Probandenausfälle zu verzeichnen. Entscheidend in diesem Zusammenhang ist die Sicherstellung der Unterstützung der Dozenten, in deren Vorlesung die zu evaluierende Lernumgebung eingesetzt wird. Je stärker der Dozent auf die Wichtigkeit der Evaluation hinweist und je mehr er die Lernumgebung in den Unterricht einbindet, desto aktiver dürften die Studierenden gewillt zeigen, sich an einer Befragung zu beteiligen oder an einem Experiment teilzunehmen.

Insgesamt erwies sich die systematische, rastergeleitete Evaluation, die verwendeten Methoden sowie die zeitliche Abstimmung der einzelnen Methoden als sehr geeignet für die Evaluation von PTO. Die feste Verankerung der Evaluationsfunktion im Entwicklerteam und das kontinuierliche Überprüfen und Testen auf verschiedenen Ebenen haben geholfen, PTO gezielt zu verbessern und zu optimieren.

4.4 Literaturverzeichnis

- Böttcher, W. (2006). Bildungsstandards und Evaluation im Paradigma der Outputsteuerung. In W. Böttcher, H. G. Holtappels & M. Brohm (Hrsg.), *Evaluation im Bildungswesen* (S. 39-49). Weinheim: Juventa.
- Kauffeld, S. (2004). *Fragebogen zur Arbeit im Team (FAT)*. Göttingen: Hogrefe.
- Kirkpatrick, D. L. (1998). *Evaluating Training Programs* (2nd ed.). San Francisco: Berrett-Koehler.
- Rowntree, D. (1992). *Exploring Open and Distance Learning*. London: Kogan Page.
- Stockmann, R. (2006). Qualitätsmanagement und Evaluation im Vergleich. In W. Böttcher, H. G. Holtappels & M. Brohm (Hrsg.), *Evaluation im Bildungswesen* (S. 23-38). Weinheim: Juventa.
- Schweibenz, W. & Thissen, F. (2003). *Qualität im Web: Benutzerfreundliche Webseiten durch Usability Evaluation*. Berlin: Springer.
- Tergan, S.O. (2000). Grundlagen der Evaluation. In P. Schenkel, S.O. Tergan & A. Lottmann (Hrsg.), *Qualitätsbeurteilung multimedialer Lern- und Informationssysteme* (S. 22–51). Nürnberg: BW Bildung und Wissen.

5.

Adaptivitätsstudie

Abstract

Die webbasierte Lernumgebung PTO (Psychopathology Taught Online) vermittelt dem Lerner Wissen zur Phänomenologie psychopathologischer Störungsbilder. Neben den multimedial aufbereiteten Lerninhalten ermöglicht PTO adaptives Lernen. Adaptives Lernen hat zum Ziel, den Lernprozess durch spezifische Lernempfehlungen zu steuern und dadurch effizientes und individuelles Lernen zu ermöglichen. Der Lerner soll nicht irgendetwas lernen, sondern sich mit denjenigen Inhalten auseinandersetzen, welche er noch am wenigsten beherrscht. PTO verfolgt hierzu einen auf Wissenskarten basierenden Ansatz, wobei der Lerner in einem ersten Schritt Ähnlichkeitsurteile zwischen je zwei Wissensobjekten abgibt. Hieraus wird mittels Nonmetrischer Multidimensionaler Skalierung (NMDS) im euklidischen Raum eine Abbildung generiert, welche die Wissensstruktur des Lerners wiedergibt. Diese Struktur wird nun mittels Prokrustes-Transformation mit einer Expertenstruktur verglichen, woraus sich über Diskrepanzen in der Struktur mangelhaftes Wissen detektieren lässt. Damit ist die Grundlage für gezielte Lernempfehlungen gegeben.

Die vorliegende Studie beinhaltet die Resultate einer Evaluationsstudie dieses in PTO implementierten Adaptiven Tutoriellen Systems. Dabei handelt es sich um ein Lernexperiment im Kontrollgruppendesign. Im Gegensatz zu vorhergehenden Untersuchungen wurde ein erstes Mal der Versuch unternommen, die Lerner statt im Labor in einem natürlichen Setting lernen lassen. Die Resultate zeigen den Lernerfolg, welcher beim adaptiv gesteuerten Lernen mit PTO erzielt werden kann. Zum einen finden sich auf rein numerischer Ebene signifikante Verbesserungen der AverageLoss-Werte, welche den Unterschied zwischen Lernern und Experten quantifizieren. Zum anderen lassen sich zwischen den verschiedenen Erhebungszeitpunkten Verbesserungen in der Wissensstruktur feststellen. Die Strukturen der Lerner nähern sich im Lernprozess systematisch derjenigen der Experten an. Weiter bestätigen sich in den Lernerkarten die positiven Effekte einer gezielten Repetition gemäss den Empfehlungen des Adaptiven Tutoriellen Systems. Insgesamt zeigen die Resultate damit die erfolgreiche Anwendung von PTO im Praxistest sowie das Funktionieren des Adaptivitätsmechanismus.

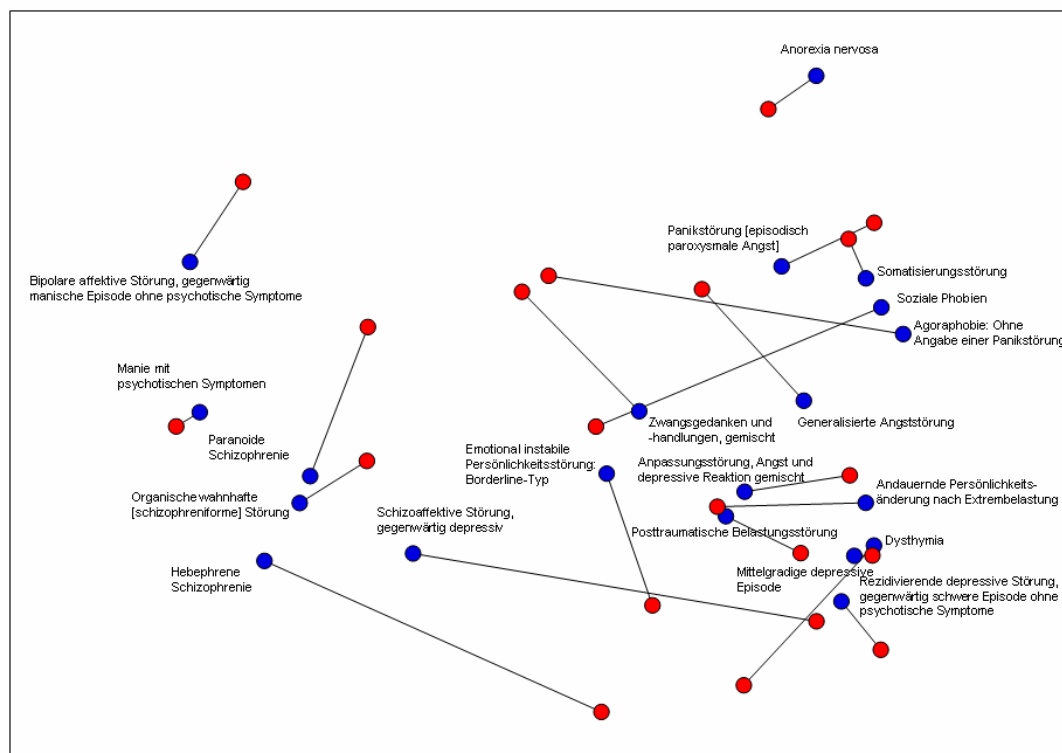
5.1 Einleitung

Die webbasierte Lernumgebung Psychopathology Taught Online (PTO) vermittelt multimedial aufbereitet Wissen zur Phänomenologie psychischer Störungen. Das didaktische Konzept von PTO ist weitgehend konstruktivistisch und basiert auf assimilativem Lernen. Dies bedeutet, dass neues Wissen an bereits bestehende individuelle Wissensstrukturen angeknüpft werden soll. Als Konsequenz daraus gliedert sich der Lernprozess von PTO in zwei Curricula. Im Curriculum 1 geht es um den Aufbau einer stabilen Grundstruktur. Dazu werden 20 der relevantesten Störungsbilder gelernt, die in der Summe das gesamte Spektrum an Störungen abbilden. Im Curriculum 2 wird diese einmal erworbene „Grundstruktur“ um 32 weitere Störungsbilder ergänzt. Das zu Grunde liegende Konzept besteht also darin, zunächst eine sachlich korrekte und kognitiv stabile Basisstruktur aufzubauen. In diese Basisstruktur kann dann weiteres Wissen eingefügt werden, ohne dass die lernende Person aufgrund neu hinzukommender, zum bisherigen Wissen eventuell inkonsistenter Information ihre gesamte Wissensstruktur umzubauen (zu akkomodieren) hat. Dieser Vorgang von assimilativem Lernen gilt in der Tradition Piagets als besonders effizient (vgl. Ausubel, 2000).

PTO zeichnet sich zudem dadurch aus, dass Wissen inhaltlich adaptiv vermittelt werden soll. Adaptives Lernen bedeutet, dass der Lernstoff in seiner Auswahl, seiner Reihenfolge und vor allem auch in der möglichen Repetition nach den momentanen Erfordernissen jedes einzelnen Individuums dargeboten werden kann. Der Lerner soll also nicht irgendetwas lernen, sondern sich mit genau jenen Inhalten auseinandersetzen, die er am wenigsten beherrscht oder gerade am besten brauchen kann.

Der dabei eingesetzte Adaptivitätsmechanismus funktioniert über drei Stufen: Möchte ein Lerner sein Wissen testen, führt er in einem ersten Schritt eine Wissensdiagnostik durch. Dabei gibt er paarweise Ähnlichkeitsurteile zwischen den gelernten Störungsbildern ab. Auf der Basis dieser Ähnlichkeitsurteile kann in einem zweiten Schritt mittels eines Skalierungsverfahrens eine grafische Abbildung des Lernerwissens in Form einer zweidimensionalen relationalen Karte generiert werden (hierbei wird der robuste NMDS-Algorithmus RobuScal verwendet; Läge, 2001; Läge, Daub, Bosia, Ryf & Jäger, 2005). Diese Kognitive Karte bildet das deklarative Merkmalswissen zu den Störungsbildern nicht direkt als lange Liste abgefragter Wissens Elemente ab, sondern indirekt als eine Form von Ähnlichkeiten und Unähnlichkeiten, die diese Person zwischen je zwei Störungsbildern wahrnimmt: Vom Lerner als ähnlich beurteilte Störungsbilder werden in der Karte nahe beieinander positioniert, während als einander unähnlich eingestufte Störungsbilder weit entfernt voneinander zu liegen kommen. Die Wissenskarte des Lerners kann nun durch eine Prokrustes-Transformation (Gower & Dijksterhuis, 2004; Läge, 2001) mit der Wissenskarte von Experten verglichen werden (vgl. hierzu Abb. 25, die aus dem weiter unten erläuterten Datensatz ein exemplarisches Beispiel für eine Kognitive

Karte und eine Prokrustes-Transformation mit dem Expertenwissen zeigt). Damit wird ein Vergleich der Wissensstruktur des Lerners mit derjenigen der Experten möglich. Das numerische Mass für diesen Vergleich ist auf der Ebene der Gesamtkarte der Average-Loss (Läge 2001). Je tiefer dieser ausfällt, desto ähnlicher sind sich die Lerner- und die Expertenstruktur. Über den Vergleich zwischen Lerner und Experten hinsichtlich jedes einzelnen Störungsbildes lassen sich Mängel in der Wissensstruktur des Lerners detektieren und fehlendes oder inadäquates Wissen aufzeigen. Technisch geschieht dies in der Prokrustes-Transformation durch den ObjectLoss, der die Unterschiedlichkeit der Positionen von zweimal demselben Objekt in Lerner- und Expertenkarte angibt. In einem dritten Schritt erhält der Lerner zur Behebung der Wissenslücken auf der Basis der Object-Losses Lernempfehlungen und spezifische Übungen zur Bearbeitung – und zwar werden hier die vier in der Wissensdiagnostik als am schlechtesten gewusst eingeschätzten Störungsbilder empfohlen. Auf diese adaptive Weise arbeitet sich eine lernende Person schrittweise im PTO-Curriculum 1 an eine gute, d.h. inhaltlich wie strukturell adäquate Wissensstruktur heran. Auf dieser Struktur aufbauend, kann sie dann in Curriculum 2 durch Vergleich des Neuen mit dem bereits Gewussten assimilativ neue Störungsbilder hinzulernen.



AvgLoss: 0.43

Abb. 25: Prokrustes-Transformation einer Lernerkarte (rote Punkte) mit der Expertenkarte (blaue Punkte). Ein AverageLoss von 0.43 spricht für eine grundsätzliche Strukturähnlichkeit zwischen der Lerner- und der Expertenkarte. Bei 4 Störungsbildern bestehen noch bedeutende Unterschiede zwischen dem Lerner und den Experten.

Die hier präsentierte Studie verfolgte insgesamt drei Ziele: Erstens soll untersucht werden, ob sich mittels Kognitiver Karten Wissen überhaupt messen lässt. Dies ist insofern von Bedeutung, als solche Karten innerhalb von PTO zur Wissensdiagnose herangezogen werden. Untersuchungen dazu wurden in der Vergangenheit bereits durchgeführt (vgl. Streule, 2007), entsprechend handelt es sich bei diesem Teil der Untersuchung um eine Replikationsstudie. Zweitens soll der Lernerfolg erfasst werden, der sich bei der Beschäftigung mit PTO erzielen lässt. Dies geschieht auf der Basis der Messung von Veränderungen in der Wissensstruktur im Lernprozess. Drittens soll die Wirkung des Adaptiven Tutoriellen Systems, welches in PTO zur Anwendung kommt, geprüft werden. Hierzu sollen auf struktureller Ebene die Veränderungen erfasst werden, die sich durch die Bearbeitung der Lernempfehlungen ergeben. Im Folgenden soll erst das methodische Vorgehen dargelegt werden. Anschliessend werden zu den einzelnen Fragestellungen die Hypothesen und die Resultate präsentiert. Anhand dreier konkreter Beispiele wird weiter der Lernprozess veranschaulicht. Zum Schluss folgt eine Diskussion der Ergebnisse.

5.2 Methode

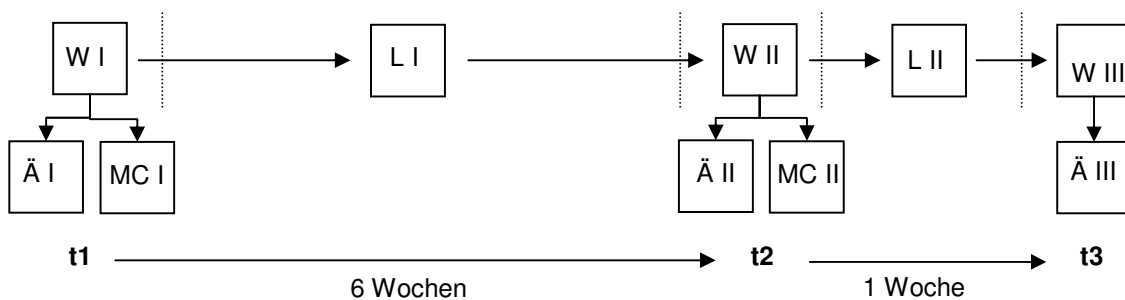
5.2.1 Experimentaldesign

Zur Beantwortung der Fragestellungen anerkennen sich die Durchführung eines Lernexperiments im Kontrollgruppendesign. PTO sollte unter möglichst realen Bedingungen genutzt und der dabei stattfindende Lernprozess erfasst werden. Der Versuchsablauf entsprach dem oben vorgestellten Lernprozess in PTO. Das Experiment fokussierte auf das Curriculum 1, bei welchem mit dem Lernen von 20 ausgewählten Störungsbildern der Aufbau einer stabilen Grundstruktur im Zentrum steht. Das Versuchsdesign für die Experimentalgruppe sah vor, dass zuerst das Wissen vor Lernbeginn erhoben werden sollte (W I). Dieser Wissenstest I bestand aus einem Multiple Choice-Test (MC I) sowie einer Erhebung mittels Ähnlichkeitsurteilen (Ä I). Der Multiple-Choice-Test diente der Überprüfung des Faktenwissens, wohingegen die Erhebung mittels Ähnlichkeitsurteilen die Erfassung des strukturellen Wissens zum Ziel hatte.

Danach folgte eine erste Lernphase (Lernphase I), in der die Probanden selbstgesteuert die 20 psychopathologischen Störungsbilder des Curriculums C1 von PTO lernten. Anschliessend erfolgte wiederum ein Wissenstest (W II). Die Wissenstests I und II fanden in einem zeitlichen Abstand (t) von 6 Wochen statt. Im Anschluss an diesen Wissenstest, der wieder aus einem Multiple-Choice-Test (MC II) sowie einer Erhebung mittels Ähnlichkeitsurteilen (Ä II) bestand, erhielten die Probanden die vier am schlechtesten gewussten Störungsbilder - detektiert aus der Ähnlichkeitserhebung II - als Lernempfehlung. Nun folgte eine zweite Lernphase (L II), die zum Ziel hatte, die vier empfohlenen Störungsbilder erneut zu lernen. Die restlichen 16 Störungsbilder wurden in dieser Phase

nicht repetiert. Für diese Lernphase stand eine Woche Zeit zur Verfügung. Nach der Repetition der vier Störungsbilder folgte abschliessend ein letzter Wissenstest (W III). Dieser bestand im Gegensatz zu den Wissenstests I und II nur aus einer Erhebung mittels Ähnlichkeitsurteilen. Abb. 26 zeigt eine Übersicht der verschiedenen Erhebungen und Lernphasen.

Experimentaldesign Lernergruppe:



W I = Wissenstest I

Ä I / II / III = Ähnlichkeitsurteile

MC I / II = Multiple-Choice-Test

L I = Lernphase I bestehend aus dem Lernen der 20 Störungsbilder des Curriculums C1

W II = Wissenstest II

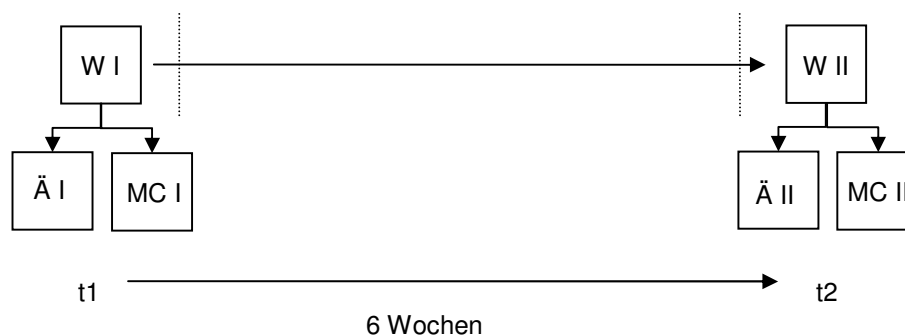
L II = Lernen der 4 am schlechtesten gewussten Störungsbilder in der Erhebung der Ähnlichkeitsurteile im Wissenstest II

W III = Wissenstest III

Abb. 26: Übersicht über den Experimental Aufbau der Lernergruppe.

Das Experimentaldesign der Kontrollgruppe sah vor, dass diese analog zur Lernergruppe die Wissenstests I und II durchführt, ohne jedoch zwischenzeitlich mit PTO zu lernen (Lernphase I entfällt somit). Der Wissenstest III wurde von der Kontrollgruppe ebenfalls nicht absolviert, da diese keine Lernempfehlungen erhielt und auch keine Repetitionsphase durchlief. Abb. 27 zeigt eine Übersicht über die Erhebungen der Kontrollgruppe.

Experimentaldesign Kontrollgruppe:



W I = Wissenstest I

Ä I / II = Ähnlichkeitsurteile

MC I / II = Multiple-Choice-Test

W II = Wissenstest II

Abb. 27: Übersicht über den Experimental Aufbau der Kontrollgruppe.

Mit dem gewählten Experimentaldesign liessen sich das Wissen zu verschiedenen Zeitpunkten im Lernprozess erheben, Wissensänderungen feststellen und damit Rückschlüsse auf den Lernprozess ziehen. Das Kontrollgruppendesign ermöglichte weiter die Kontrolle von Effekten, welche sich durch die reine Bearbeitung der Wissenstests ergaben.

5.2.2 Wissenstests

Die folgenden Ausführungen beziehen sich in erster Linie auf die Lernergruppe. Wie oben erläutert durchlief die Kontrollgruppe nur die Wissenstests I und II.

Der Wissenstest I diente der Erfassung des Wissensstands vor Lernbeginn und bestand aus einem Multiple-Choice-Test sowie der auf PTO integrierten Wissensdiagnostik, bei der das Wissen über paarweise Ähnlichkeitsurteile zwischen je zwei Störungsbildern erfasst wurde (insgesamt 190 Ähnlichkeitsurteile). Aus den Ähnlichkeitsdaten wurde mittels Skalierungsverfahren - dem NMDS-Algorithmus RobuScal - eine zweidimensionale Struktur errechnet, welche die Wissensstruktur der Probanden zu den 20 Störungsbildern des Curriculums 1 repräsentierte. Diese wurde wiederum mit einer Expertenstruktur verglichen und dabei auf Abweichungen von derselben geprüft. Dadurch liess sich mangelhaftes Wissen (falsch oder nicht gewusste Störungsbilder) detektieren.

Der Multiple-Choice-Test setzte sich aus insgesamt 40 Fragen (je 2 Fragen zu einem der 20 Störungsbilder) zusammen. Die Fragen bezogen sich inhaltlich auf die Lerninhalte

von PTO und wurden den Probanden jeweils in randomisierter Reihenfolge präsentiert. Für die Berechnung des Punktescores ergab ein richtig gesetztes Kreuz einen Punkt. Falsch gesetzte Kreuze gaben einen Punkt Abzug.

Der Wissenstest II diente der Erhebung des Wissensstands nach der Lernphase. Dabei wurden wiederum ein Multiple Choice-Test mit 40 weiteren Fragen (erneut 2 Fragen zu jedem der 20 Störungsbilder) sowie die ähnlichkeitsbasierte Wissensdiagnostik eingesetzt. Um einen systematischen Einfluss des nicht völlig identischen Schwierigkeitsniveaus der Fragen der beiden Multiple-Choice-Tests auf die Resultate zu verhindern, wurde die Reihenfolge, in der die beiden Fragebogen präsentiert wurden, für je eine Hälfte der Lernergruppe unterschiedlich festgelegt (dasselbe galt auch für die Kontrollgruppe). Die maximale Punktzahl im Test A lag bei 70, im Test B bei 72 Punkten. Die aus den Ähnlichkeitsurteilen resultierende Kognitive Karte wurde nun wiederum mit einer Expertenstruktur verglichen. Die vier am schlechtesten gewussten Störungsbilder konnten dadurch bestimmt und zur Repetition empfohlen werden.

Der Wissenstest III hatte zum Ziel, die Wissensstruktur nach Repetition der vier am schlechtesten gewussten Störungsbilder zu erfassen und bestand nur noch aus der Erhebung der Ähnlichkeitsurteile. Die Kognitive Karte, die aus diesen Daten generiert werden konnte, stellte die finale Wissensstruktur des Experiments dar. Vergleiche mit den Karten der Wissenstests I und II erlaubten daher für jeden Probanden der Lernergruppe die Rückschlüsse auf den Lernprozess. Tab. 6 zeigt einen tabellarischen Überblick über die im Verlauf des Experiments erhobenen Daten.

Tab. 6: Überblick über die erhobenen Daten in den Wissenstests (in chronologischer Abfolge des Testverlaufs).

| Ä I (Ähnlichkeitsurteile) | MC I (Fragebogen) | Ä II (Ähnlichkeitsurteile) | MC II (Fragebogen) | Ä III (Ähnlichkeitsurteile) |
|--|--|--|--|--|
| 190 Urteile zwischen allen 20 Störungsbildern. | 40 Fragen (jeweils 2 Fragen pro Störungsbild). | 190 Urteile zwischen allen 20 Störungsbildern. | 40 Fragen (jeweils 2 Fragen pro Störungsbild). | 190 Urteile zwischen allen 20 Störungsbildern. |

Insgesamt führte die Experimentalgruppe damit 3 Ähnlichkeitserhebungen und zwei Multiple-Choice-Tests durch. Abgesehen von der Ähnlichkeitserhebung III erfolgte die Bearbeitung der Tests durch die Kontrollgruppe analog.

5.2.3 Lernmaterial und Lernsituation

Als Lernmaterial dienten die 20 Störungsbilder des Curriculums C1 von PTO. Diese wurden während der Lernphase I von den Probanden gelernt. Die 4 Störungsbilder, die in Lernphase II gelernt wurden, bildeten jeweils ebenfalls eine Auswahl aus den 20 Stö-

rungsbildern des Curriculums C1 (die vier am schlechtesten gewussten). PTO bietet zum Lernen eines Störungsbilds allgemeine Informationen zu Phänomenologie und Verlauf, interaktiv aufbereitete Lerninhalte, Fallbeispiele, aber auch Selbsttests, Lernziel-Checks und Übungen. Sämtliche Materialien standen den Probanden während der Lernphasen zur freien Verfügung.

Alle für das Experiment relevanten Aktivitäten konnten die Probanden von einem beliebigen Rechner mit Internetanschluss erledigen. Dadurch sollte den Probanden die Möglichkeit gegeben werden, in ihrem gewohnten und natürlichen Umfeld mit PTO zu lernen. Bewusst wurde auf eine künstliche Lernumgebung (z.B. Lernlabor) verzichtet, da die Lernsituation derjenigen der zukünftigen Nutzer von PTO entsprechen sollte.

Bezüglich der Lernphase I wurden die Probanden instruiert, die 20 Störungsbilder des Curriculums C1 samt Kategorieeinführung prüfungsreif zu lernen. Um Kontrolle über die effektive Lernzeit wahren zu können, wurden die Probanden angehalten, jede Lerneinheit in einem Lernprotokoll festzuhalten und dabei anzugeben, wieviel Zeit sie in das Lernen eines bestimmten Störungsbilds investiert hatten. Für die Lernphase II lautete die Instruktion, die 4 vom Adaptiven Tutoriellen System zur Repetition empfohlenen Störungsbilder erneut zu lernen.

5.2.4 Testpersonenkollektiv

Für das Lernexperiment waren Probanden nötig, die dem Profil der zukünftigen Nutzer der Lernumgebung entsprachen. Da PTO in erster Linie für die universitäre Ausbildung im Bereich Psychologie konzipiert wurde, schienen Studierende der Psychologie diese Anforderung am besten zu erfüllen. Die Probanden sollten zudem bereits Vorwissen im Bereich der Psychopathologie aufweisen, da die Erhebung des Wissens mittels Ähnlichkeitsurteilen nur zu sinnvollen Resultaten führt, wenn bereits ein Mindestmass an Vorwissen vorhanden ist (wichtig für die Ähnlichkeitserhebung des Wissenstests I). Weiter sollten die Probanden – zumindest im Falle der Lernergruppe - gewillt sein, sich längere Zeit mit PTO zu beschäftigen und sich die Lerninhalte anzueignen.

Die Personen der Kontrollgruppe sollten sich zudem bezüglich des Vorwissens auf demselben Niveau wie die Lernergruppe bewegen. Insbesondere im Hinblick auf die quantitative Auswertung der Ergebnisse sollte die Stichprobe über eine bestimmte Grösse verfügen, welche die geplanten statistischen Analysen erlaubt. Insgesamt wurden 46 Personen rekrutiert. 30 Personen für die Lernergruppe und weitere 16 Personen für die Kontrollgruppe. Wie bei einem Lernexperiment diesen Ausmasses, welches ein beträchtliches Mass an Disziplin und Selbstorganisation seitens der Probanden verlangt, zu erwarten, gab es Ausfälle während der Testphase. Schliesslich absolvierten 12 Personen in der Experimental-, respektive weitere 12 Personen in der Kontrollgruppe die Wissenstests I und II. Der Wissenstest III wurde noch von 7 Personen der Lernergruppe durchgeführt. Das

Durchschnittsalter der Probanden lag bei 28 Jahren (Standardabweichung 7.7). 17 Probandinnen waren weiblichen und 7 Probanden männlichen Geschlechts.

5.3 Fragestellung 1: Wissensmessung mit Kognitiven Karten

Die erste Fragestellung lautete, ob sich mittels Kognitiver Karten Wissen messen lässt.

5.3.1 Hypothese

Wie erläutert, erfolgt die Wissensdiagnostik in PTO anhand von Kognitiven Karten. Das Erfassen von Wissensstrukturen und insbesondere die Diagnose von Wissen anhand Kognitiver Karten ist ein bislang wenig verbreitetes Verfahren. In verschiedenen Studien konnte bereits gezeigt werden, dass sich Wissen mittels Kognitiver Karten messen lässt (vgl. Streule, 2007). Die Experimentalsettings dieser Untersuchungen waren eher künstlicher Natur. Der vorliegenden Untersuchung liegt hingegen erstmalig ein natürliches Setting zugrunde.

Über die Multiple-Choice-Tests liess sich das Faktenwissen der Probanden zu den Störungsbildern erfassen. Die Punktzahl im Test entspricht damit der Qualität des Faktenwissens (die grundlegende Annahme dabei besagt: je mehr Punkte, desto qualitativ besser ist das Wissen). Mit dem Vergleich der Kognitiven Karte eines Probanden mit der Expertenkarte liess sich hingegen die Güte der Wissensstruktur überprüfen. Dabei gilt, dass je kleiner der AverageLoss ausfällt, die Wissensstruktur des Probanden umso besser ist. Für die vorliegende Untersuchung ist nun der Zusammenhang zwischen dem Abschneiden im Multiple-Choice-Test und der Güte der Wissensstruktur eines jeden Probanden entscheidend. Die Resultate im Multiple Choice-Test fungieren dabei als externes Kriterium. Vorausgesetzt mit den Kognitiven Karten lässt sich tatsächlich Wissen messen, dann müsste sich zwischen den Kognitiven Karten und dem externen Validierungsinstrument (Multiple Choice-Test) ein systematischer Zusammenhang zeigen (Postuliert wird ein negativer Zusammenhang, da der AverageLoss ein Abweichungsmass darstellt). Die Hypothese lautet damit:

Hypothese:

- Je kleiner der AverageLoss, desto besser ist das Wissen (verglichen mit Faktenwissen erhoben über einen Multiple Choice-Test).

5.3.2 Resultate

Zur Prüfung der Hypothese wurde die Korrelation der Resultate von Wissensdiagnostik (AverageLoss-Werte) und Multiple-Choice-Test zu den beiden Zeitpunkten t1 und t2 untersucht. Tab. 7 zeigt die Korrelation zwischen der Anzahl Punkte im Multiple-Choice-Test und den AverageLoss-Werten für die Lerner- als auch für die Kontrollgruppe zu den Zeitpunkten t1 und t2. Zudem wurde die Korrelation über alle Probanden zu beiden Zeitpunkten berechnet.

Tab. 7: Korrelation zwischen Punktzahl im Multiple-Choice-Test und AverageLoss-Werten aus der Wissensdiagnostik.

| Zeitpunkt | Lernergruppe | Kontrollgruppe |
|------------------|---|--|
| t1 | 0.18 Nicht signifikant; $p=28.5\%$ | -0.35 Nicht signifikant; $p=12,9\%$ |
| t2 | -0.59 <i>Signifikant; $p=2.1\%$</i> | -0.46 Nicht signifikant; $p=6.4\%$ |
| Beide Zeitpunkte | -0.39 <i>Signifikant; $p=0.3\%$</i> | |

Für den Zeitpunkt t1 fällt weder die Korrelation der Lernergruppe noch diejenige der Kontrollgruppe signifikant aus. Dies lässt sich damit begründen, dass wenn nur schon in einem der beiden Tests schlechte Resultate erzielt werden, keine Korrelation zustande kommt. Da weder die Probanden der Lerner- noch der Kontrollgruppe sich zu diesem Zeitpunkt mit den Inhalten des Curriculums 1 auseinandergesetzt hatten, war dieses Resultat zu erwarten. Dasselbe gilt für die Kontrollgruppe zum Zeitpunkt t2. Hier fällt die Korrelation mit -0.46 jedoch bereits ziemlich hoch aus und wird nur knapp nicht signifikant. Dieses Resultat gibt Hinweis darauf, dass sich nur schon rein durch das Durchführen der Wissensdiagnostik, respektive durch die Bearbeitung des Multiple-Choice-Tests leichte Verbesserungen der Wissensstruktur erzielen lassen. Bei der Lernergruppe zeigt sich zum Zeitpunkt t2 ein anderes Bild. Hier fällt die Korrelation mit -0.59 sehr hoch aus. Dieses Resultat erweist sich zudem auf dem 5%-Niveau als signifikant.

Über beide Zeitpunkte und alle Testpersonen gerechnet ergibt sich ebenfalls eine signifikante Korrelation von -0.39 (siehe Abb. 28). Im mittleren Punktebereich des Multiple-Choice-Tests streuen die AverageLoss-Werte stark. Dagegen zeigt sich, dass im hohen Punktebereich keiner der Probanden einen hohen AverageLoss aufweist. Weiter erzielten die Probanden, welche im Multiple-Choice-Test eine tiefe Punktzahl erreichten, in den allermeisten Fällen auch hohe Werte beim AverageLoss.

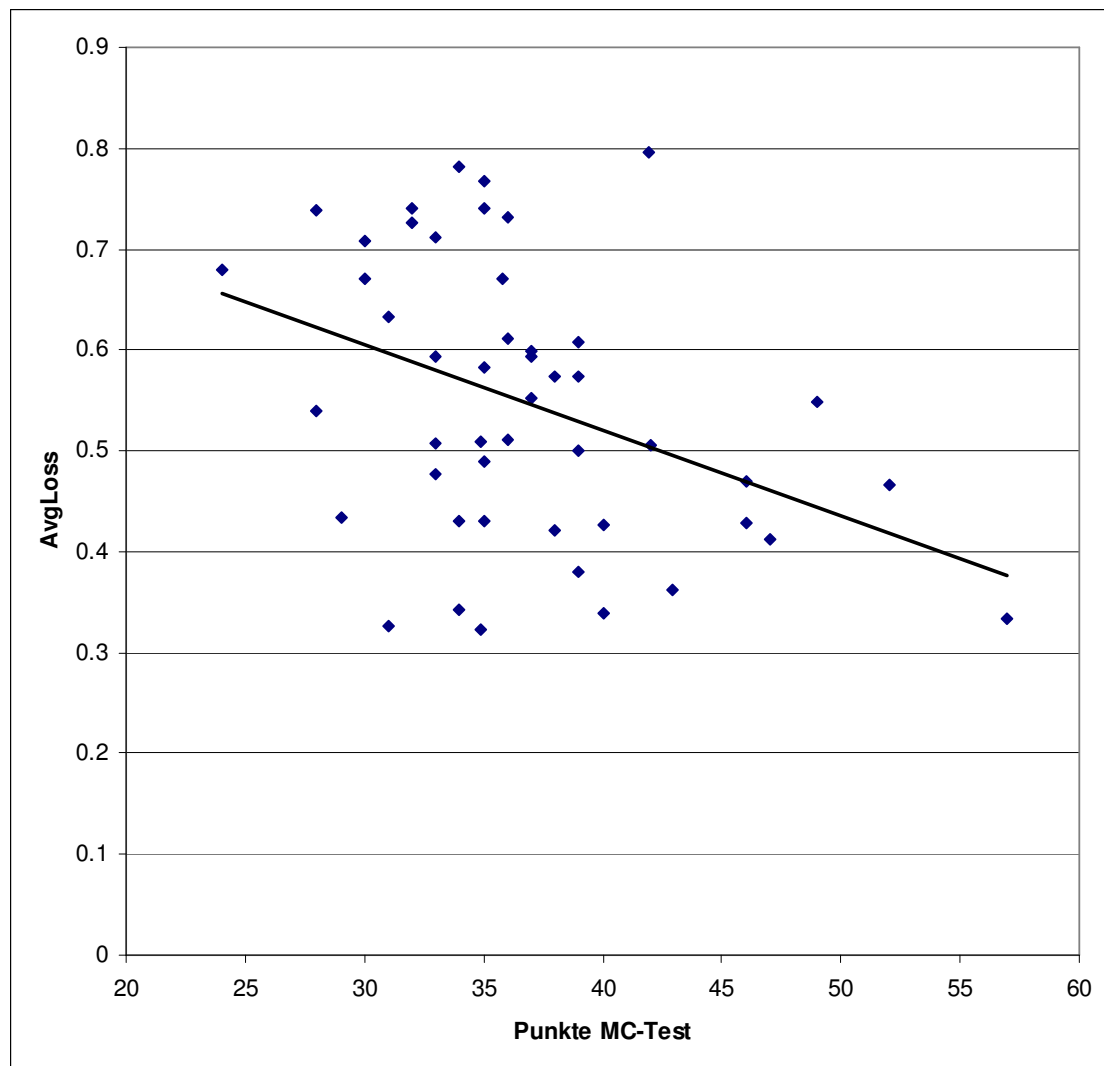


Abb. 28: Zusammenhang zwischen AverageLoss und Punktzahl im Multiple-Choice-Test für beide Zeitpunkte und über alle Probanden gerechnet.

Die hohe Korrelation von -0.59 für die Lernergruppe zeigt einen starken Zusammenhang zwischen den Ergebnissen in der Wissensdiagnostik und denjenigen im Multiple-Choice-Test, welcher das externe Kriterium darstellt. Dieses Resultat bestätigt damit, dass sich mittels Kognitiver Karten Wissen messen lässt. Dieses Ergebnis, welches bereits in anderen Untersuchungen (siehe oben) gefunden wurde, konnte damit auch für natürliche Untersuchungssettings bestätigt werden.

5.4 Fragestellung 2: Erfassung des Lernerfolgs

Die zweite Fragestellung befasste sich mit der Überprüfung des Lernerfolgs, der sich bei der Arbeit mit PTO erzielen lässt.

5.4.1 Hypothesen

Das Ziel dieser Untersuchung bestand darin, den Lernfortschritt zu erfassen, der beim Lernen mit PTO erzielt werden kann. Es sollte geprüft werden, ob diese Lernfortschritte messbar sind und wie sich das Wissen über die psychopathologischen Störungsbilder im Lauf des Lernprozesses ändert. Die Lernfortschritte sollten dabei zum einen auf der Ebene des Faktenwissens und zum anderen auf der Ebene des strukturellen Wissens erhoben und dokumentiert werden. Die Erfassung des Faktenwissens erfolgte über einen Multiple-Choice-Test, wohingegen das strukturelle Wissen über ein ähnlichkeitsbasiertes Verfahren erhoben wurde. Von zentralem Interesse waren die Veränderungen in den Wissenskarten, die die Struktur des Wissens veranschaulichen. Die Frage war, ob sich Veränderungen des Wissens im Lauf des Lernprozesses in den Karten nachweisen lassen. Innerhalb von PTO besteht das grundlegende Ziel darin, die Wissensstruktur der Lerner durch den Lernprozess einer Expertenstruktur anzugleichen. Die quantitative Grösse für diesen Strukturunterschied ist der AverageLoss. Je ähnlicher sich die Lerner- und die Expertenstruktur sind, desto kleiner wird der AverageLoss. Somit wurde erwartet, dass der AverageLoss im Lauf des Lernprozesses sinkt. Weiter wurde postuliert, dass sich durch die Beschäftigung mit PTO auch das Faktenwissen verbessert und die Probanden damit im Multiple-Choice-Test nach der Lernphase besser abschneiden würden. Damit ergaben sich die folgenden Hypothesen:

Hypothese 1:

- Bei den Lernern nimmt der AverageLoss durch den Lernprozess von t1 zu t2 und weiter zu t3 hin ab.

Hypothese 2:

- Bei den Lernern nimmt die Punktzahl im Multiple-Choice-Test von t1 zu t2 zu.

5.4.2 Resultate

Die AverageLoss-Werte nehmen von t1 (vor Lernbeginn) nach t2 (nach Lernphase I) signifikant ab ($p=4,0\%$; $d=0.79$) (siehe Abb. 29). Da bei t3 (nach der Repetitionsphase) nur

die Daten von 7 Personen einfließen, ist eine inferenzstatistische Auswertung des Unterschieds nicht möglich.

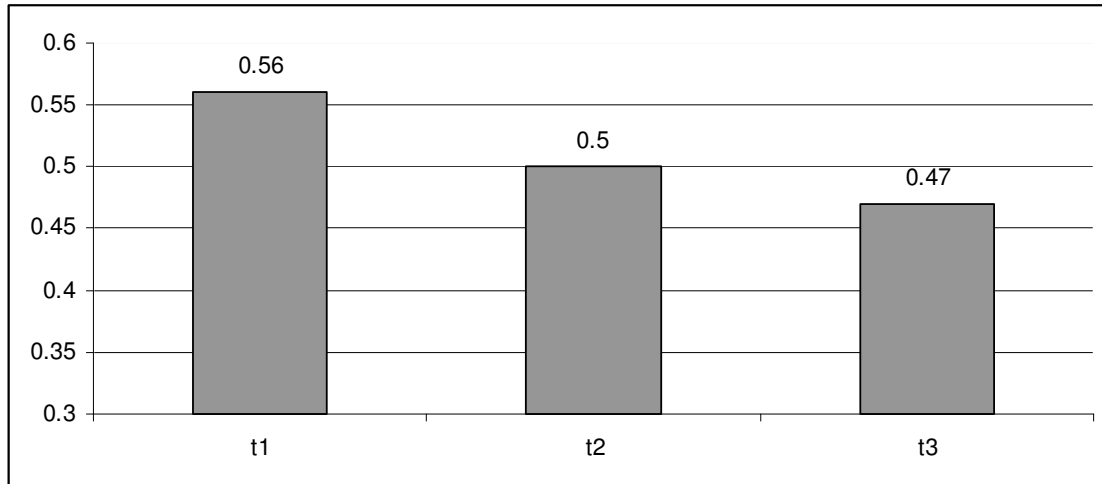


Abb. 29: Durchschnittliche AverageLoss-Werte für die Erhebungszeitpunkte t1, t2 und t3.

Die signifikante Abnahme des AverageLoss-Wertes zeigt, dass sich die Wissensstruktur der Probanden bezüglich der in der Lernphase gelernten psychopathologischen Störungsbilder verbessert und sich der Expertenstruktur annähert. Im Multiple-Choice-Test schneiden die Probanden der Lernergruppen nach der Lernphase ebenfalls signifikant besser ab ($p=0.09\%$; $d=2.03$) (siehe Abb. 30).

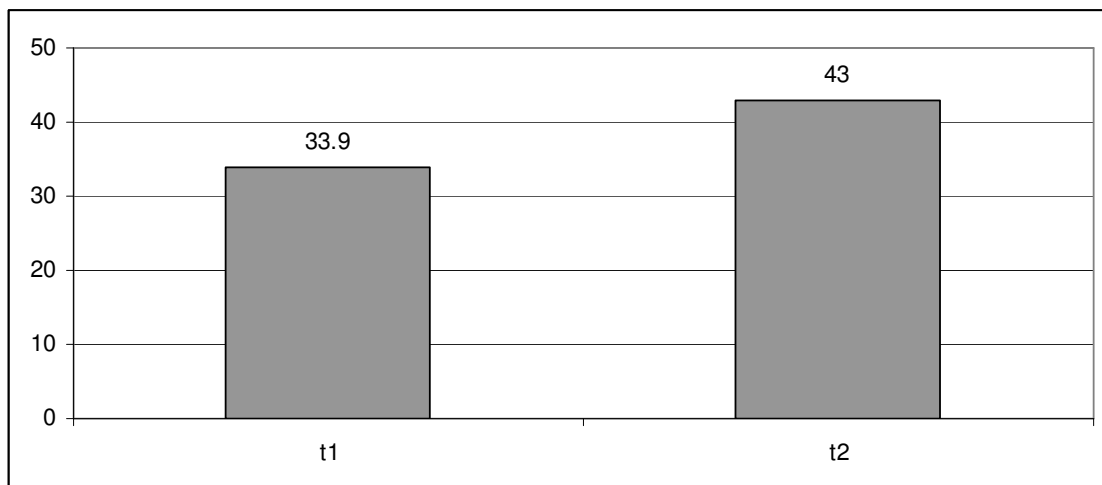


Abb. 30: Durchschnittliche Punktzahl im Multiple-Choice-Test für die Erhebungszeitpunkte t1 und t2.

Die Resultate zeigen, dass sich die Probanden durch die Auseinandersetzung mit den Lerninhalten von PTO sowohl auf der Ebene des Faktenwissens als auch auf der Ebene der Wissensstruktur signifikant verbessert haben: Mit zunehmendem Wissen über die Störungsbilder hat sich auch das relationale Gefüge – die Wissensstruktur bezüglich der Störungsbilder – deutlich verbessert.

Die Veränderung der Wissensstruktur der Probanden lässt sich mittels einer „Loss-oriented Meta Map“ kurz LOMM-Karte (Läge et al., 2005; bei Läge, 2001 noch als „prokrustes-basierte Personenskalierung“ bezeichnet) veranschaulichen. In einer solchen Karte sind die einzelnen Testpersonen durch jeweils einen Punkt repräsentiert. Jene Personen, die eine ähnliche Wissensstruktur aufweisen, werden in der Karte auch nahe beieinander positioniert. Die Karte zeigt die Positionen der Kontrollgruppe (Zeitpunkte t1 und t2) sowie diejenigen der Lerngruppe (Zeitpunkte t1, t2 und t3). Die Expertenstruktur wird durch einen weiteren Punkt repräsentiert (siehe Abb. 31).

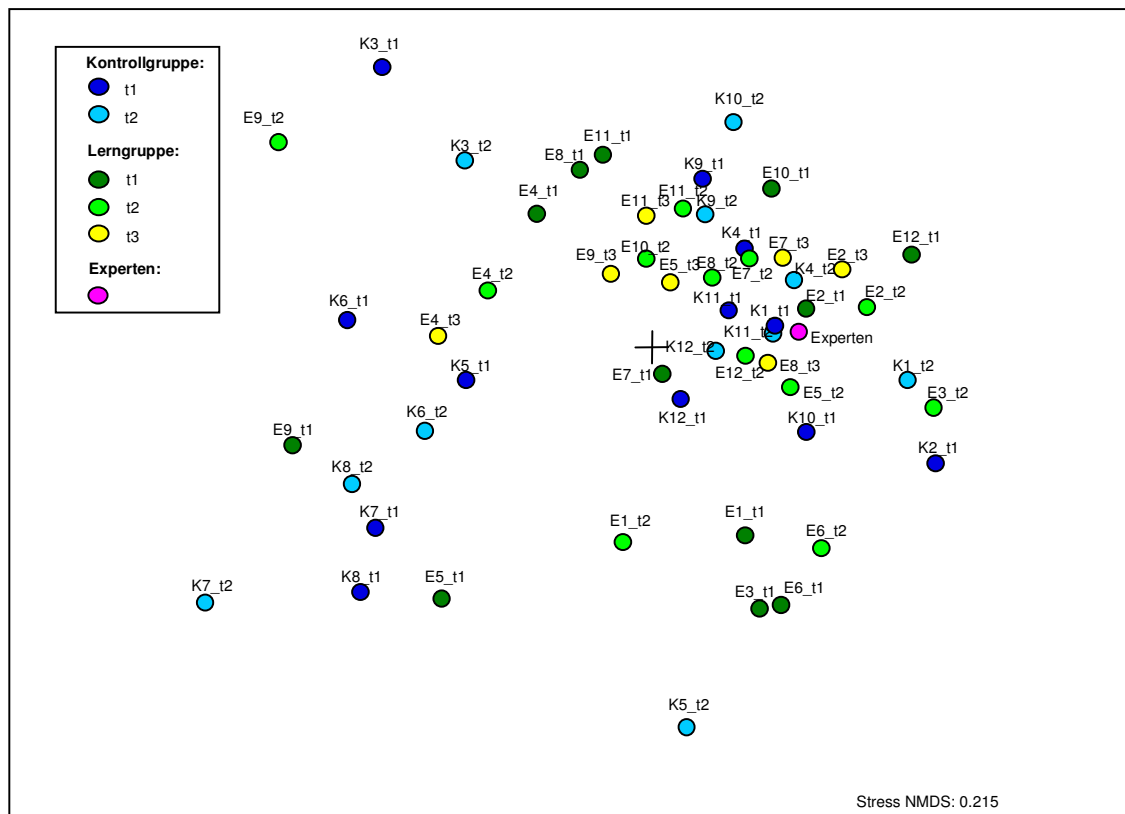


Abb. 31: LOMM-Karte mit Kontrollgruppe und Lerngruppe zu den verschiedenen Erhebungszeitpunkten.

Mit dieser LOMM-Karte lässt sich die Wissensstruktur der einzelnen Probanden zu den verschiedenen Erhebungszeitpunkten miteinander in Beziehung setzen und vergleichen.

Da sich die Expertenposition ebenfalls als ein räumlich definierter Punkt in die Karte einbeziehen lässt, zeigt sich, welche Probanden nahe bei der Expertenposition zu liegen kommen, also eine Struktur aufweisen, die derjenigen der Experten ähnlich ist.

Die Karte weist stark darauf hin, dass die Personen, die nicht oder noch nicht mit PTO gelernt haben, in der Peripherie zu liegen kommen und allgemein die grösste Distanz zur Expertenposition aufweisen. Dabei handelt es sich um Personen der Kontrollgruppe (t1 und t2) sowie Personen der Experimentalgruppe, bevor diese zu lernen begonnen hatten (t1). Die Wissensstruktur dieser Personen weicht damit am stärksten von derjenigen der Experten ab. Je weiter fortgeschritten die Probanden im Lernprozess sind, desto stärker nähern sie sich der Expertenposition an. So kommen die Positionen der Lerner nach der Lernphase (t2) allgemein näher bei der Expertenposition zu liegen. Nach der Repetitionsphase (t3) verstärkt sich dieser Effekt nochmals. In der LOMM-Karte ist damit die klare Tendenz erkennbar, dass sich durch das Lernen mit PTO die Wissensstruktur des Lerners derjenigen eines Experten annähert.

Die durchschnittliche Distanz zwischen der Lerner- und der Expertenposition zu den Zeitpunkten t1 und t2 lässt sich quantifizieren (siehe Abb. 32). Die Distanz verringert sich deutlich, jedoch knapp nicht in signifikantem Ausmass ($p=5.1\%$; $d=0.79$). Zum Zeitpunkt t3 beträgt die durchschnittliche Distanz noch 0.74. Bei der Kontrollgruppe nimmt die Distanz dagegen vom Zeitpunkt t1 (1.22) zum Zeitpunkt t2 (1.28) hin in nicht signifikantem Ausmass ($p=39.8\%$; $d=0.07$) zu.

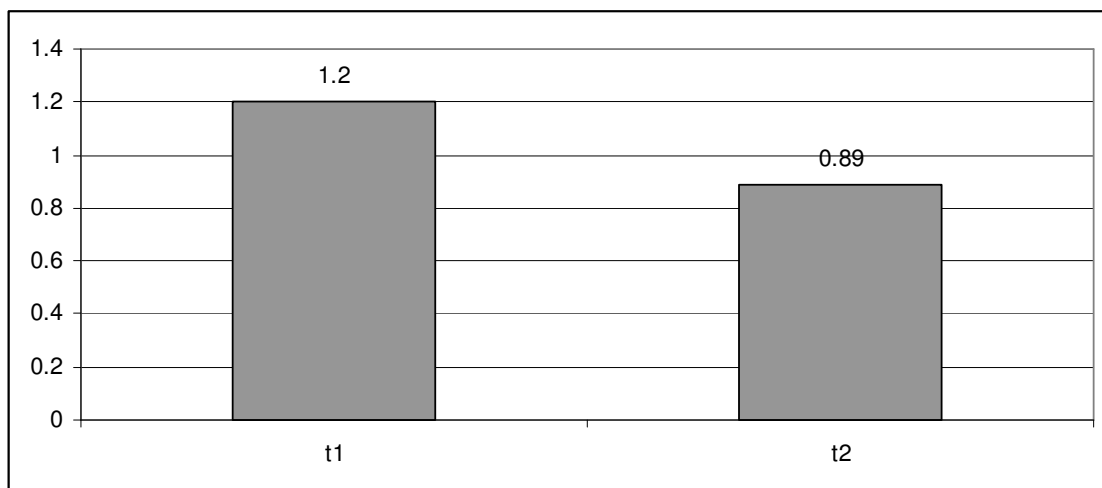


Abb. 32: Durchschnittliche Distanz zwischen Lerner- und Expertenposition zu den Zeitpunkten t1 und t2 (Massstab: durchschnittliche Distanz aller Punkte zum Kartenschwerpunkt = 1).

Es zeigt sich die starke Tendenz, dass sich die Probanden durch die Lernphase generell der Expertenposition annähern und sich damit in Richtung Expertenstruktur verbessern.

Im Folgenden sollen die Probanden der Lernergruppe gesondert betrachtet werden (siehe Abb. 33). Die Cluster markieren die verschiedenen Erhebungszeitpunkte.

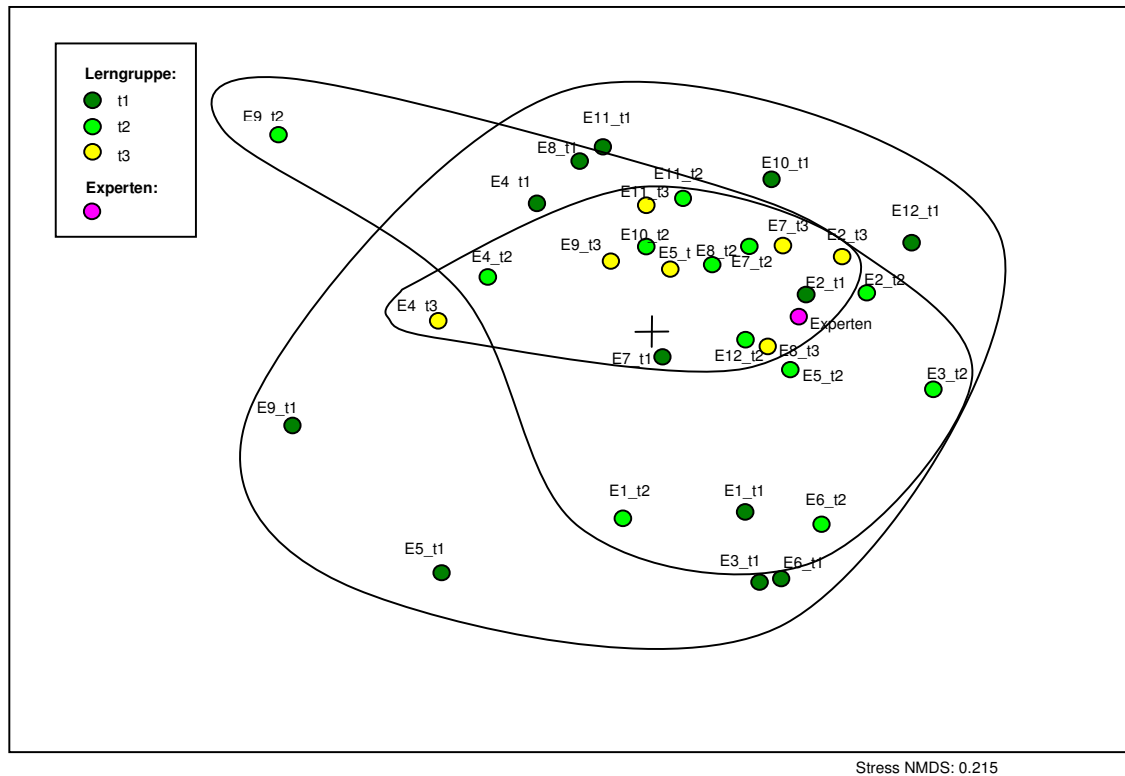


Abb. 33: LOMM-Karte mit den Positionen der Lernergruppen. Die Cluster kennzeichnen die drei Erhebungszeitpunkte t1, t2 und t3.

Die Karte zeigt, dass die Cluster über den Lernprozess hinweg enger werden und sich stärker zur Expertenposition hin zusammenziehen. Der Cluster t2 weist mit der Person E9 einen markanten Ausreisser auf. Gesamthaft betrachtet nähern sich durch das Lernen mit PTO die Wissensstrukturen der Lernergruppe sukzessive derjenigen der Experten an (zu beachten ist jedoch, dass die Stichprobe zum Zeitpunkt t3 5 Personen weniger umfasst als zum Zeitpunkt t2).

Dieselbe Karte lässt sich auch mit den Testpersonen der Kontrollgruppe rechnen. Die Cluster markieren hierbei die Daten der Erhebungszeitpunkte t1 und t2 (siehe Abb. 34).

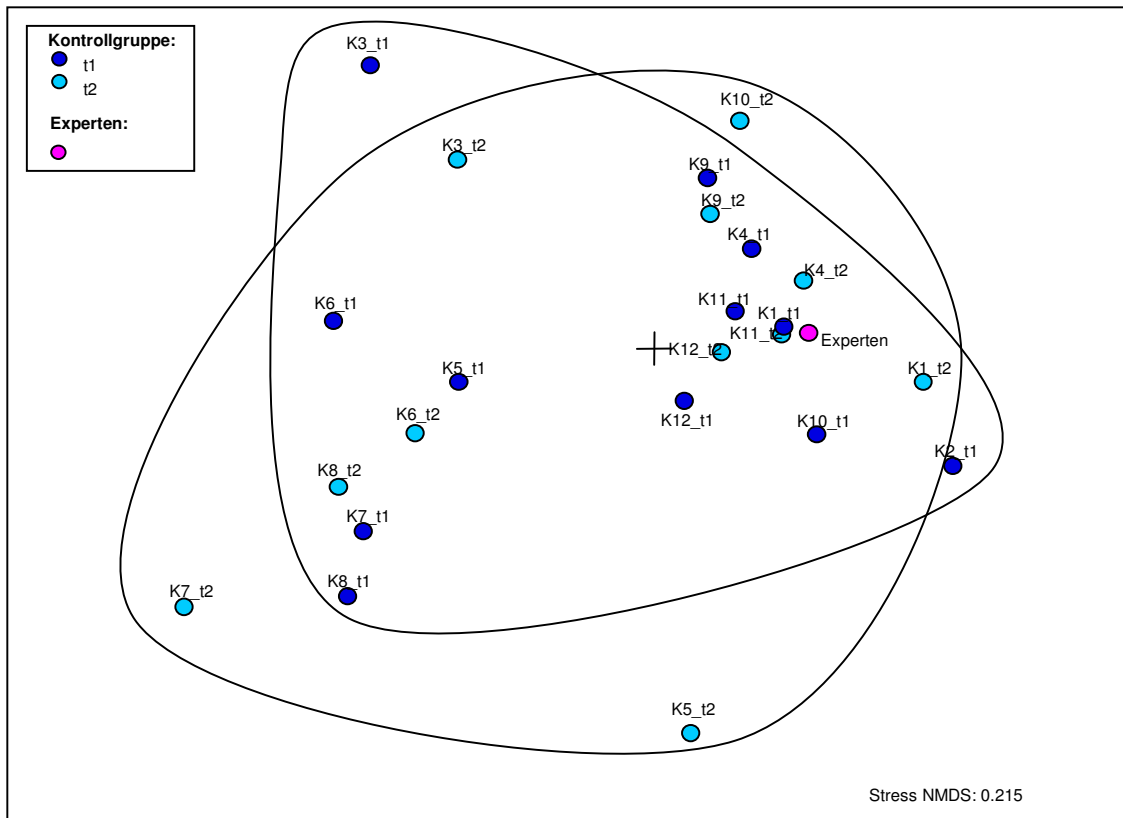


Abb. 34: LOMM-Karte mit den Positionen der Kontrollgruppe. Die Cluster markieren die Erhebungszeitpunkte t1 und t2.

Im Gegensatz zur Karte mit der Lernergruppe sind die beiden Cluster der Kontrollgruppe fast gleich gross. Die Probanden der Kontrollgruppe nähern sich von t1 zu t2 hin also nicht der Expertenposition an. Die Wissensstruktur verändert sich im Lauf der Zeit nur minim und wird qualitativ nicht besser.

5.5 Fragestellung 3: Wirkung des Adaptiven Tutoriellen Systems

Die dritte Fragestellung behandelte die Wirkung des Adaptiven Tutoriellen Systems, welches in PTO in Form der Wissensdiagnostik angeboten wird.

5.5.1 Hypothesen

Das Adaptive Tutorielle System von PTO hat das Ziel zu messen, über welche Wissensstruktur jemand verfügt, um dann dort mit spezifischen Lernempfehlungen ansetzen zu können, wo das Wissen des Lernalers am wenigsten korrekt ist.

Bei der vorliegenden Untersuchung stand die Frage nach dem Effekt des Adaptiven Tutoriellen Systems im Zentrum. Hat der Lerner das Curriculum C1 von PTO bearbeitet, ver-

fügt er über ein gutes Basiswissen zu 20 ausgewählten psychopathologischen Störungsbildern. In der Wissensdiagnostik kann er nun sein Wissen testen, respektive seine Wissensstruktur erheben und diese mit der Struktur eines Experten vergleichen. Die vier im Vergleich zu den Experten am schlechtesten gewussten Störungsbilder werden in der Folge als Lernempfehlungen zur Repetition empfohlen. Die Untersuchung sollte zeigen, wie sich das Repetieren dieser Störungsbilder auswirkt. Dies zum einen auf Gesamtebene betrachtet über die durchschnittliche Verbesserung im Vergleich zu den Experten als auch auf individueller Ebene im Bezug auf die Wissensstruktur eines einzelnen Lerner. Damit ergeben sich die folgenden Hypothesen:

Hypothese 1:

- Bei einer Prokrustes-Transformation zwischen Lernerkarte und Expertenkarte verringern die repetierten Störungsbilder ihre ObjectLoss-Werte (verglichen zum Zeitpunkt vor der Repetition).

Hypothese 2:

- Bei einer Projektion der zu repetierenden Störungsbilder in die Expertenkarte führt die Repetition zu einer Annäherung der projizierten Positionen an die korrespondierenden Störungsbilder in der Expertenkarte.

5.5.2 Resultate

Um die Annäherung der repetierten Störungsbilder an die Expertenpositionen zu untersuchen, wurde die Änderung der durchschnittlichen Distanz in der Kognitiven Karte untersucht. Verglichen wurden dabei die ObjectLoss-Werte vor und nach der Repetition pro Störungsbild. Der ObjectLoss ist das Mass für die Abweichung zwischen Experten- und Lernerposition auf Objektebene (der ObjectLoss ist das analoge Mass zum AverageLoss auf Objektebene). Bei 7 Probanden, die die Erhebungen zu den Zeitpunkten t2 und t3 absolvierten, entspricht dies 28 ObjectLoss-Werten (4 repetierte Störungsbilder von 7 Probanden) pro Erhebungszeitpunkt (siehe Abb. 35). Die Überprüfung erfolgte mittels T-Test für abhängige Stichproben.

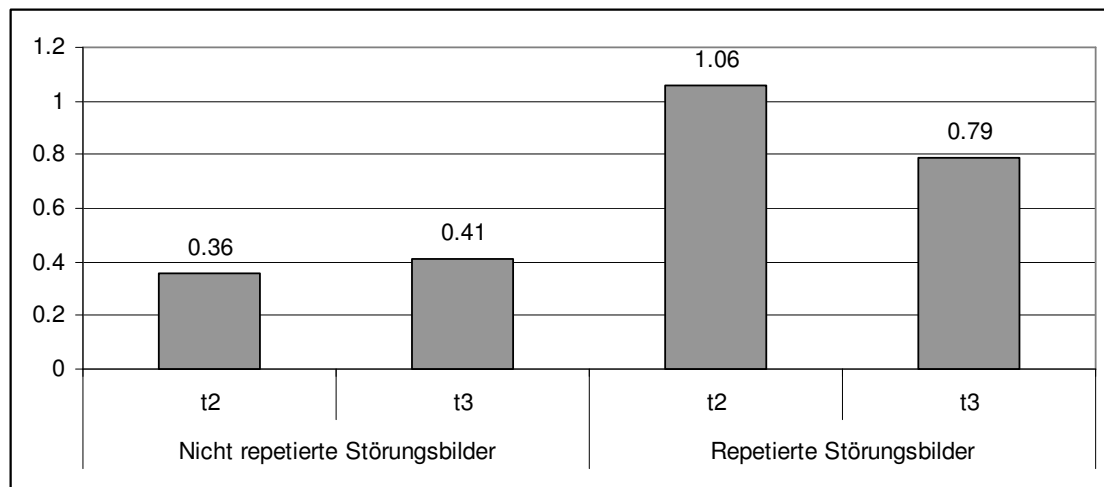


Abb. 35: Veränderung der ObjectLoss-Werte von t2 nach t3 für repetierte und nicht repetierte Störungsbilder.

Die durchschnittlichen ObjectLoss-Werte der 28 repetierten Störungsbilder nehmen von 1.06 auf 0.79 signifikant ab ($p=0.1\%$; $d=0.87$). Die Werte der nicht repetierten Störungsbilder steigen hingegen signifikant von 0.36 auf 0.41 ($p=2.6\%$; $d=0.26$). Da es sich bei den untersuchten Werten um Extremwerte handelte, muss beachtet werden, dass die Resultate durch eine allgemeine Tendenz zur Mitte beeinflusst sein können. Die Resultate zeigen, dass sich bei den nicht repetierten Störungsbildern die ObjectLoss-Werte erwartungsgemäss leicht verschlechtern, während sich die erneute Beschäftigung mit den am schlechtesten gewussten Störungsbildern positiv auf die Wissensstruktur auswirkt.

Auf individueller Ebene lässt sich weiter untersuchen, wie die repetierten Störungsbilder ihre Position innerhalb der Kognitiven Karte ändern. Dazu lassen sich in die Expertenkarte die 4 repetierten Störungsbilder eines einzelnen Probanden rechnen. Dies lässt sich für die beiden Erhebungszeitpunkte t2 und t3 in derselben Karte durchführen. Dadurch werden die Änderungen der Positionen der repetierten Objekte in Bezug zur Expertenposition sichtbar.

Die beiden folgenden positiven Beispiele der Versuchspersonen E5 und E9 zeigen deutlich, dass sich die repetierten Störungsbilder von t2 nach t3 Richtung Expertenposition bewegen (Abb. 36 und Abb. 37). Bei der Versuchsperson E5 bestehen keine grossen Unterschiede im Ausmass, in dem die 4 Objekte die Position ändern. Die Verbesserungen fallen für alle repetierten Störungsbilder ähnlich stark aus.

Bei der Testperson E9 korrigieren die Objekte 10 und 12 stärker hin zur Expertenposition, wobei auch für die Objekte 5 und 19 Verbesserungen erkennbar sind. Die beiden Testpersonen unterscheiden sich aber bezüglich dem Grad, mit dem es ihnen gelungen ist, die repetierten Störungsbilder dem Expertenniveau anzunähern. Dieses Ziel hat die

Testperson E5 deutlich besser erreicht. Ihre Störungsbilder positionieren sich zum Zeitpunkt t3 in der Nähe der Expertenposition. Bei der Testperson E9 kann zwar eine Veränderung in Richtung Expertenposition festgestellt werden, jedoch bleibt insbesondere für die Objekte 5 und 10 selbst nach der Repetition noch deutliches Verbesserungspotential.

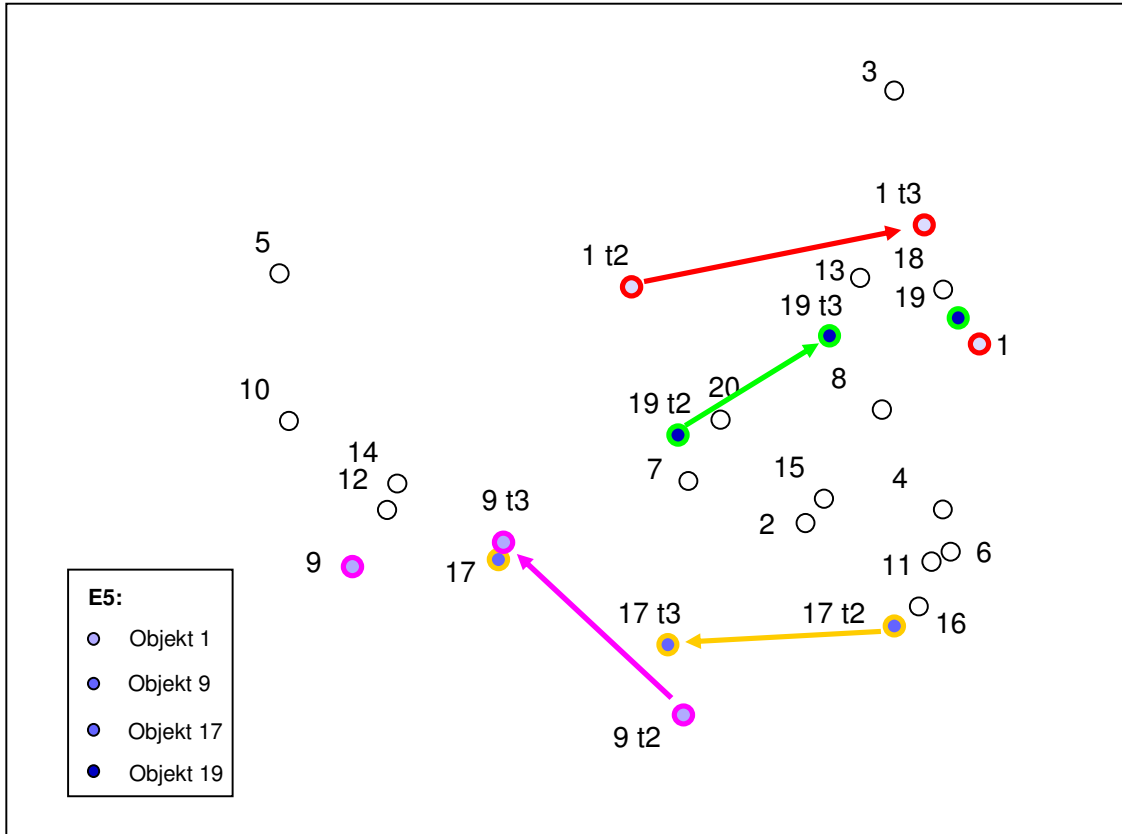


Abb. 36: Positionsänderungen von t2 nach t3 der 4 repetierten Objekte in Bezug zu den Expertenpositionen der Versuchsperson E5.

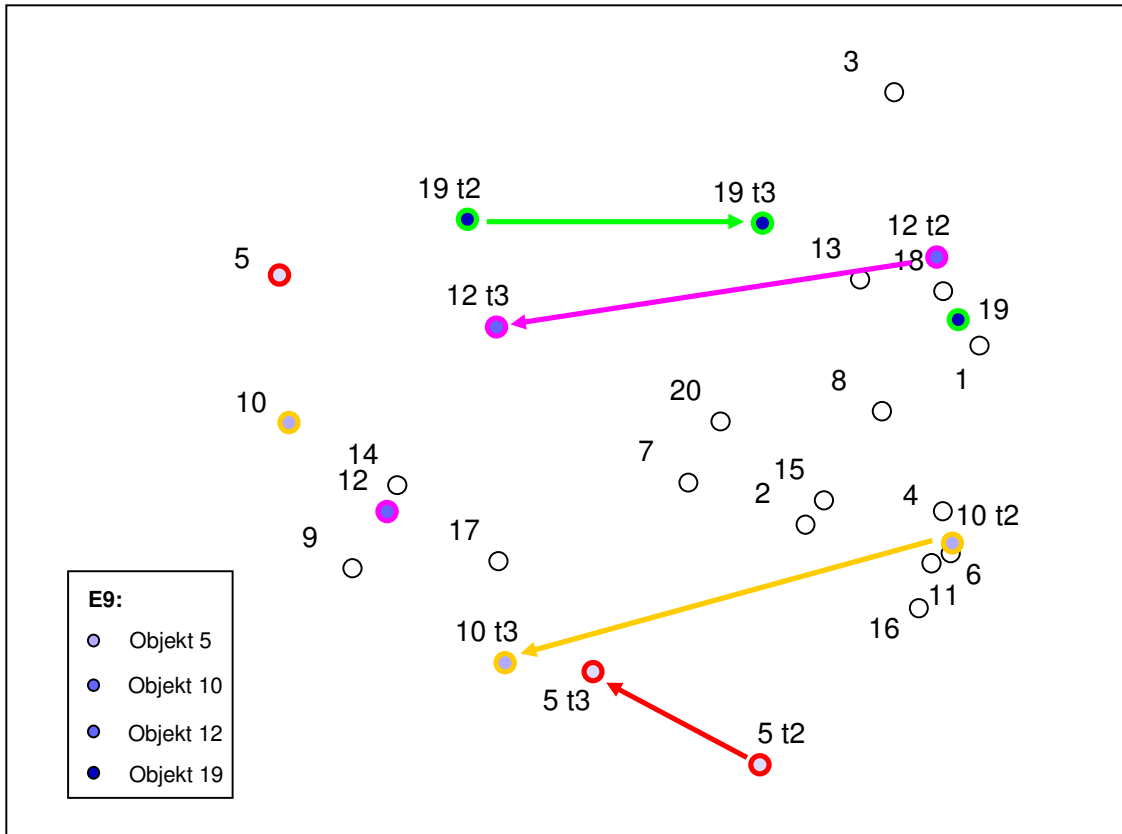


Abb. 37: Positionsänderungen von t2 nach t3 der 4 repetierten Objekte in Bezug zu den Expertenpositionen der Versuchsperson E9.

Die Resultate zeigen, dass es den beiden Versuchspersonen gelungen ist, durch die Repetition ihr Wissen zu den zum Zeitpunkt t2 am schlechtesten gewussten Störungsbilder zu verbessern. Entsprechende Veränderungen sind in den beiden Karten klar erkennbar und bestätigen auf individueller Ebene das Wirken des in PTO implementierten Adaptivitätsmechanismus.

Bei der Versuchsperson E7 handelt es sich um das Beispiel einer Person, der es nicht gelungen ist, durch die Repetition eine qualitative Verbesserung der Kartenstruktur zu erzielen (siehe Abb. 38). Das Störungsbild 1 nähert sich zwar der Expertenposition an. Für das Störungsbild 20 ist dies in geringem Ausmass ebenfalls noch der Fall. Dagegen bleibt die Position des Objekts 17 stabil und verschlechtert sich sogar deutlich im Fall des Objekts 18. Trotz Repetition hat es diese Probandin insgesamt nicht geschafft, ihre Wissensstruktur zu verbessern. Die repetitierten Objekte verändern sich in ihrer Position aber nur geringfügig, was für die Stabilität der Struktur und die Reliabilität der Messung spricht.

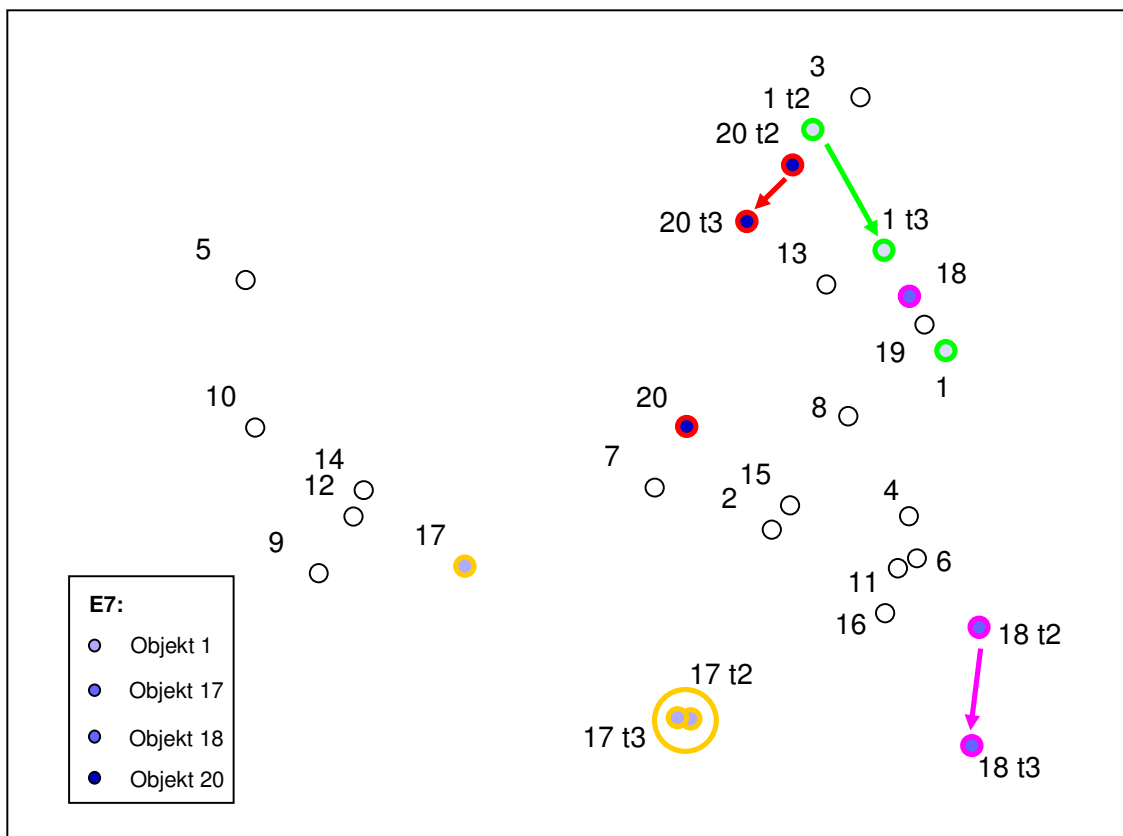


Abb. 38: Positionsänderungen von t2 nach t3 der 4 repetierten Objekte in Bezug zu den Expertenpositionen der Versuchsperson E7.

Der Versuchsperson E2 ist es kaum gelungen, eines ihrer 4 repetierten Störungsbilder merklich in Richtung der Expertenposition zu verbessern (Abb. 39). Sie verfügt jedoch bereits zum Zeitpunkt t2 über eine ausserordentlich gute Wissensstruktur (AverageLoss-Wert von 0.33) und zeigt diese auch zum Zeitpunkt t3 (AverageLoss-Wert von 0.38).

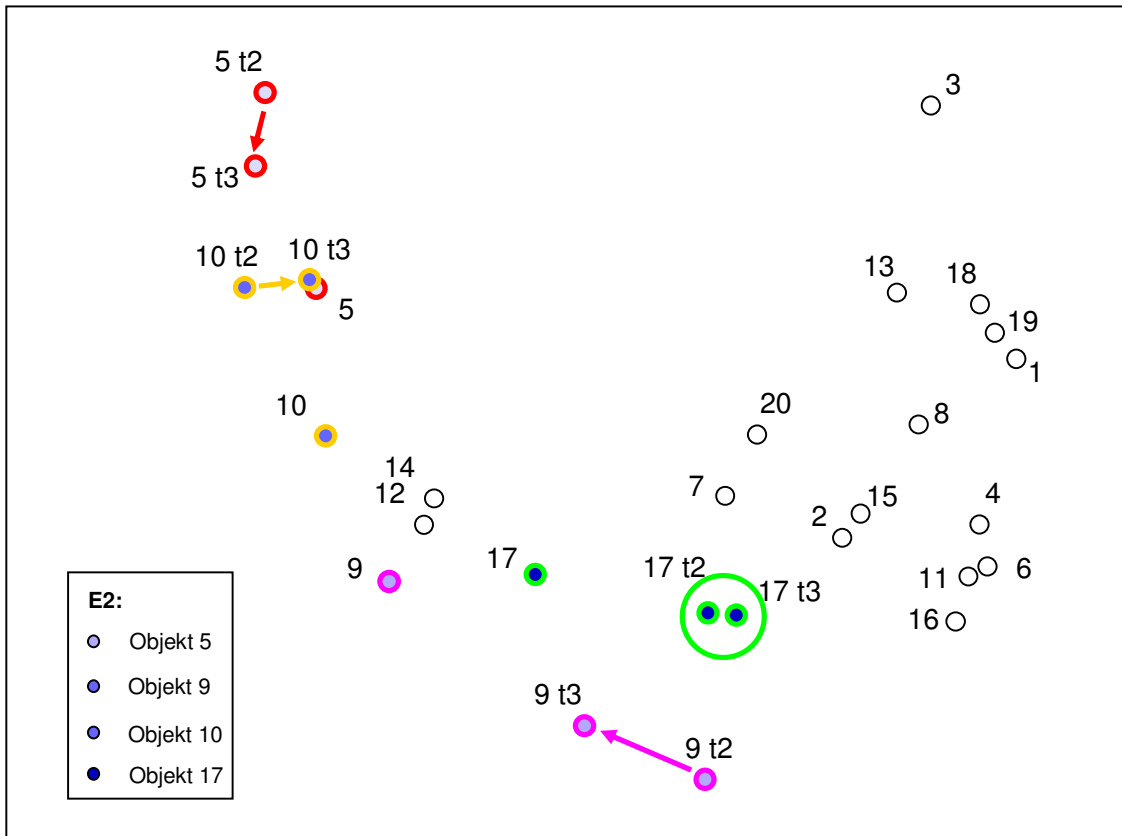


Abb. 39: Positionsänderungen von t2 nach t3 der 4 repetierten Objekte in Bezug zu den Expertenpositionen der Versuchsperson E2.

Diese Stabilität der Wissensstruktur lässt sich für diese Testperson auch in der LOMM-Karte untersuchen. Wie Abb. 40 zeigt, ändert sich ihre Position innerhalb der Karte nur minim. Daraus lässt sich schliessen, dass die Probandin E2 über eine sehr stabile Wissensstruktur zu den Störungsbildern verfügt, die der Expertenstruktur sehr nahe kommt und die auch durch die Repetition nicht massgeblich beeinflusst wird.

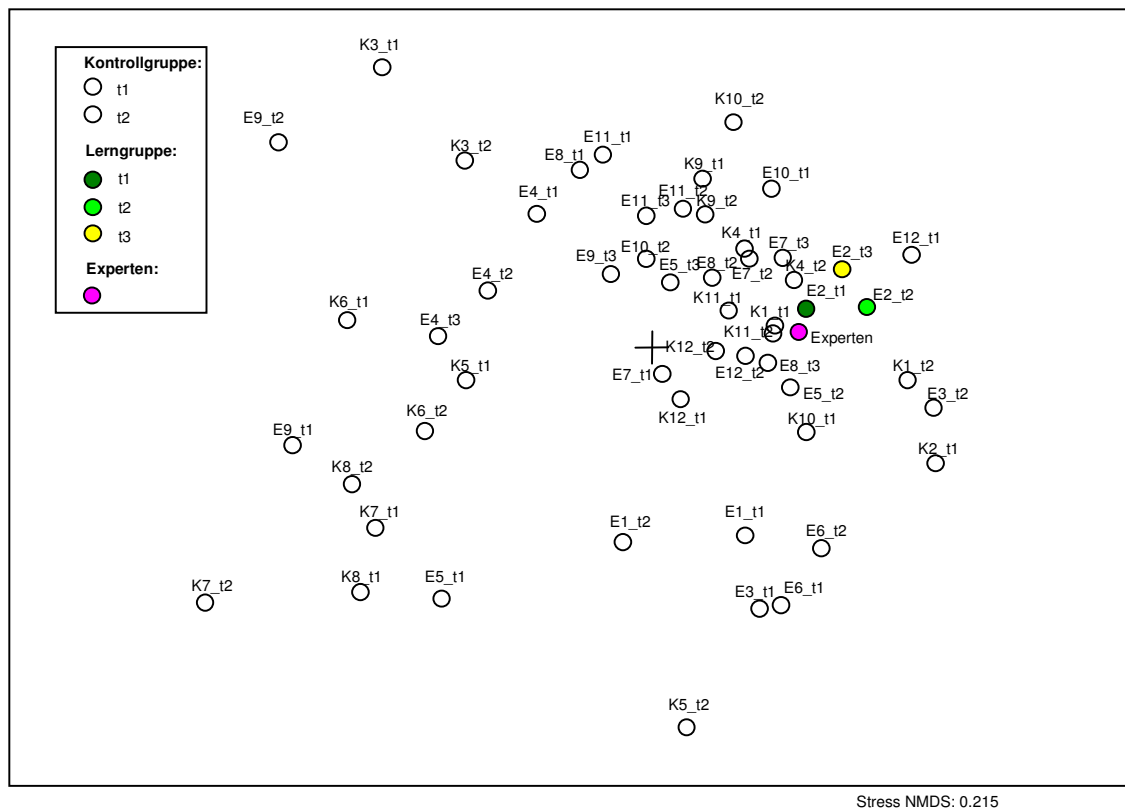


Abb. 40: Position der Versuchsperson E2 vor Lernbeginn (t1), vor der Repetitionsphase (t2) und nach der Repetitionsphase (t3).

Auf inferenzstatistischer Ebene lässt sich zur Frage nach der Verbesserung des Average-Loss-Wertes von t2 (vor Repetitionsphase) nach t3 (nach Repetitionsphase) aus Mangel an Daten keine Aussage machen. Die Untersuchung der Änderung der ObjectLoss-Werte zeigt zwar eine signifikante Veränderung für denselben Zeitraum. Jedoch muss hier die Tendenz zur Mitte berücksichtigt werden, welche bei solchen Extremwerten zum Tragen kommt. Auf individueller Ebene lässt sich, wie am Beispiel der Versuchsperson E2 demonstriert, das Funktionieren des Adaptivitätsmechanismus nachweisen.

5.6 Lernprozess auf Individualebene

Im Folgenden soll anhand dreier konkreter Beispiele der Lernprozess und die Generierung einer Wissensstruktur beim Lernen mit PTO veranschaulicht werden. Dabei wird bewusst auf ein positives, ein durchschnittliches wie auch auf ein negatives Beispiel zurückgegriffen, um unterschiedliche Entwicklungen im Prozess der Wissensentwicklung zu illustrieren. Abb. 41 zeigt die Veränderung der Wissensstruktur der Versuchsperson E8 über die Erhebungszeitpunkte t1, t2 und t3. Dabei handelt es sich um ein sehr positives Beispiel. Vor Lernbeginn befindet sich die Versuchsperson E8 in der Peripherie der LOMM-Karte. Durch die Bearbeitung der Lektionen zu den 20 Störungsbildern (Lernphase I) gelingt es dieser Person zum einen, sich Faktenwissen anzueignen, was sich in einer Steigerung im Multiple-Choice-Test (von 39 auf 52 Punkte) niederschlägt. Zum andern verbessert sich ihre Wissensstruktur durch den Lernprozess stark. Sie kommt zum Erhebungszeitpunkt t2 deutlich näher bei der Expertenposition zu liegen. Durch die Repetition der 4 am schlechtesten gewussten Störungsbilder nähert sich ihre Wissensstruktur weiter derjenigen der Experten an (Erhebungszeitpunkt t3). Diese Versuchsperson hat somit durch die Arbeit mit PTO auf der Ebene des Merkmalswissen als auch auf der Ebene des relationalen Wissens grosse Fortschritte verzeichnet.

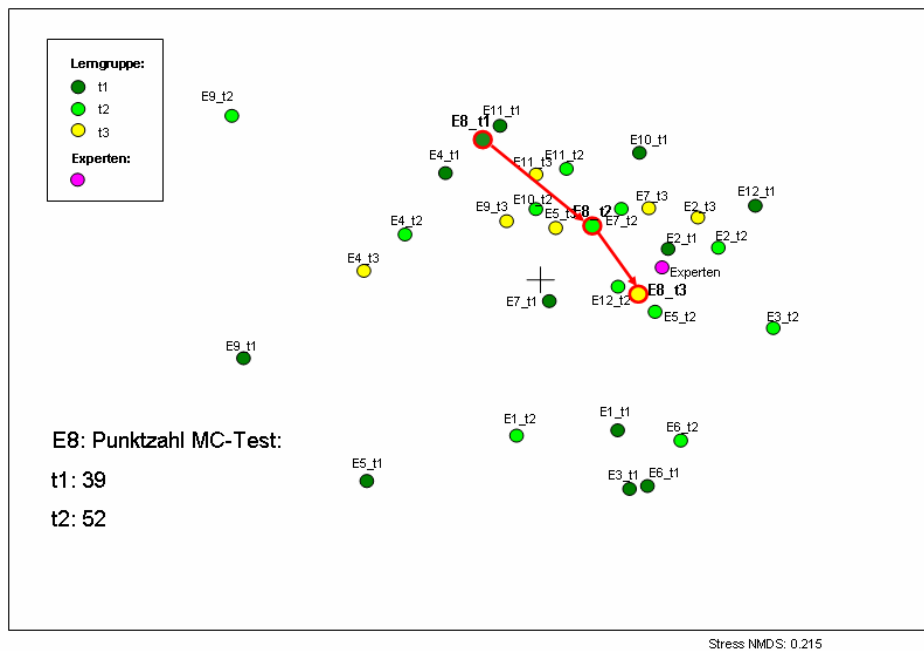


Abb. 41: Positives Beispiel: Entwicklung der Wissensstruktur der Versuchsperson E8.

Am Beispiel der Versuchsperson E11 zeigt sich ein durchschnittlicher Lernverlauf. Sie verbessert sich von t1 nach t2 von einer peripheren Lage in Richtung Expertenposition. Dies spiegelt sich auch auf der Ebene des Faktenwissens, wo sich diese Person von 33

auf 46 Punkte zu verbessern vermag. Die Repetition der 4 am schlechtesten gewussten Störungsbilder führt dann aber nicht zu einer qualitativen Verbesserung des relationalen Wissens. In der Wissenskarte bewegt sich die Versuchsperson nur leicht und nähert sich dabei nicht der Expertenposition an (siehe Abb. 42).

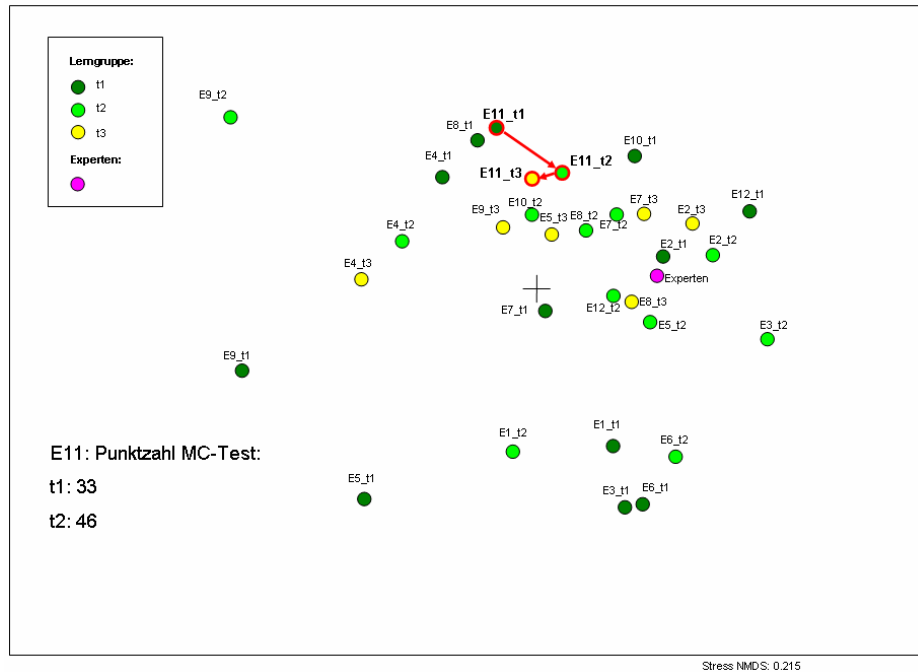


Abb. 42: Durchschnittliches Beispiel: Entwicklung der Wissensstruktur der Versuchsperson E11.

Ein eher negatives Beispiel stellt die Versuchsperson E4 dar (siehe Abb. 43). Diese verbessert sich zwar im Multiple-Choice-Test von 35 auf 49 Punkte, entwickelt jedoch keine korrekte Wissensstruktur. Im Gegenteil entfernt sie sich in der Karte von t1 nach t2 zu t3 hin sukzessive von der Expertenposition. Zudem bewegt sie sich allgemein von der Struktur anderer Probanden weg. Sie ändert auch nicht zufällig ihre Position in der Karte, sondern platziert sich ziemlich konsistent im selben Bereich. Diese Person scheint über eine idiosynkratische - falsche aber konsistente - Struktur zu verfügen.

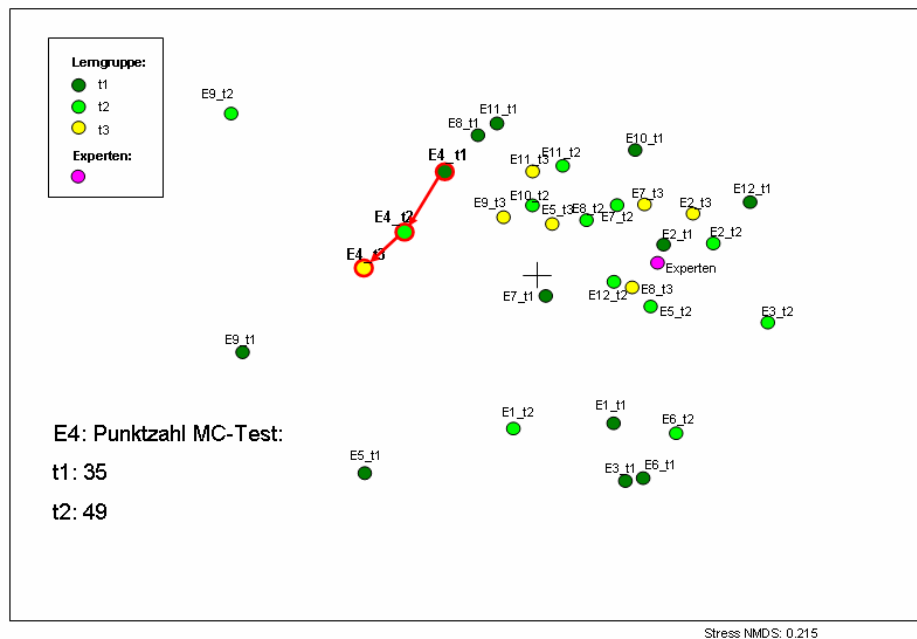


Abb. 43: Negatives Beispiel: Entwicklung der Wissensstruktur der Versuchsperson E4.

Die drei Beispiele auf individueller Ebene veranschaulichen auf plastische Weise den Lernprozess und insbesondere die Änderung der Wissensstruktur. Es zeigt sich, dass ein besseres Faktenwissen nicht zwingend mit einem besseren relationalen Wissen einhergeht. Korrektes Wissen zu einzelnen Störungsbildern ist ein notwendiges, aber kein hinreichendes Kriterium für eine korrekte Wissensstruktur, da hier die korrekte relationale Verankerung der Störungsbilder entscheidend ist.

5.7 Diskussion

Die Lernumgebung PTO verfügt mit der Wissensdiagnostik über ein komplexes Verfahren, welches dem Lerner helfen soll, seinen Lernprozess effizient und effektiv zu gestalten. Dieses Adaptive Tutorielle System wurde mit der vorliegenden Studie unter realen Bedingungen einer eingehenden Prüfung unterzogen. Die Studie verfolgte dabei dreierlei Fragestellungen: Erstens sollte geprüft werden, ob sich Wissen mittels Kognitiver Karten überhaupt messen lässt. Da das Adaptive Tutorielle System auf diesen Wissenskarten basiert, ist dies eine grundlegende Voraussetzung für den Einsatz eines solchen Systems. Zweitens sollte der Lernerfolg erfasst werden, der sich bei der Arbeit mit PTO erzielen lässt. Drittens sollte der Adaptivitätsmechanismus überprüft werden, welcher in PTO zur Anwendung kommt.

Mit der Studie konnte erstmals für eine natürliche Lernsituation der Nachweis erbracht werden, dass sich Wissen mittels Kognitiver Karten messen lässt. Der Wissenszuwachs, gemessen über das externe Kriterium Resultat im Multiple-Choice-Test korrelierte signi-

fikant mit dem AverageLoss, dem Mass für die Güte der Wissensstruktur. Der Zusammenhang zwischen der Güte einer Wissensstruktur in einer Karte und dem Wissen gemessen über einen Multiple-Choice-Test konnte nachgewiesen werden. Dies bedeutet, dass damit die Grundlage für das Funktionieren des Adaptiven Tutoriellen Systems gegeben ist und dass solche Kognitive Karten ein valables Instrument zur Wissensmessung darstellen.

Weiter liess sich der Lernerfolg auf der Ebene des deklarativen Wissens als auch auf der Ebene der Wissensstrukturen nachweisen. Die Verbesserung des deklarativen Wissens belegen die signifikant besseren Resultate im Multiple-Choice-Test nach der Bearbeitung von PTO. Zum andern verbesserte sich aber auch das relationale Wissen über die Störungsbilder, wie die nach der Lernphase signifikant tieferen AverageLoss-Werte zeigen. Die Arbeit mit PTO führt damit zu Verbesserungen des deklarativen Wissens als auch zu qualitativ besseren Wissensstrukturen.

Die Wissensstrukturen der Versuchspersonen verbesserten sich generell und näherten sich dem Expertenniveau an. Die Resultate zeigen sich dabei nicht nur auf numerischer Ebene, sondern zeichnen sich auch in den individuellen Kognitiven Karten ab. Die Karten veranschaulichen deutlich die Fortschritte einzelner Lerner auf der Strukturebene. Der Vergleich der eigenen Wissensstruktur mit derjenigen der Experten gibt dem Lerner die einzigartige Möglichkeit, sein Wissen über das reine Faktenwissen hinaus auf einer strukturellen Ebene zu prüfen. Somit bestätigte sich mit dem durchgeführten Lernexperiment auf quantitativ numerischer - als auch rein visuell in den Karten erkennbar – das Funktionieren des in PTO verwendeten Adaptiven Tutoriellen Systems. Aus den Resultaten lässt sich weiter ableiten, dass der Einsatz eines so aufgebauten Adaptiven Tutoriellen Systems in grundsätzlich allen merkmalsbasierten Wissensbereichen denkbar ist und sich nicht auf den Bereich psychopathologischer Störungsbilder beschränken muss.

5.8 Literaturverzeichnis

- Ausubel, D. P. (2000). *The acquisition and retention of knowledge*. Dodrecht: Kluwer Academic Publishers.
- Gower, J. C. & Dijksterhuis, G. B. (2004). *Procrustes problems*. Oxford: Oxford University Press.
- Läge, D. (2001). *Ähnlichkeitsbasierte Diagnostik von Sachwissen*. Habilitationsschrift an der Philosophischen Fakultät der Universität Zürich.
- Läge, D., Daub, S. Bosia, L., Jäger, C. & Ryf, S. (2005). *Die Behandlung ausreißer-behafteter Datensätze in der Nonmetrischen Multidimensionalen Skalierung – Relevanz, Problemanalyse und Lösungsvorschlag*. AKZ-Forschungs-bericht Nr. 21. Zürich: Angewandte Kognitionspsychologie.
- Streule, R. (2007). *Sag' mir, was ich noch nicht weiss - Individualisierung des Wissenserwerbs über adaptive Auswahl der Lerninhalte mittels Nonmetrischer Multidimensionaler Skalkierung und Prokrustes-Transformation*. Zürich: Zentralstelle der Studentenschaft.

6.

**Studie zur Lerneffizienz
und Lerneffektivität**

Abstract:

Die technischen Möglichkeiten im Bereich E-Learning eröffnen neue Perspektiven in der Präsentation von Lehrinhalten. Interaktivität und dynamische Abbildungen können bei geschicktem und didaktisch begründetem Einsatz Lernprozesse unterstützen. Für die vorliegende Studie wurden in der virtuellen Lernumgebung "PTO – Psychopathology Taught Online" drei Funktionalitäten auf ihre Lernwirksamkeit hin überprüft: eine interaktive, Text und Bild verknüpfende Grafik, eine Texthinterleuchtungsfunktion zur Hervorhebung der relevantesten Stellen und ein in den Fliesstext integriertes Pop-up-Glossar. Die Studie zeigt, dass sich die Implementierung solcher Techniken in einer deutlich gesteigerten Lerneffizienz (messbar an kürzeren Lernzeiten bei gleichem oder besserem Lernerfolg) niederschlägt. Gleichzeitig lässt sich die Lerneffektivität (im Sinne von Vollständigkeit und korrekter Systematisierung) bei kognitiv anspruchsvollen Inhalten steigern.

6.1 Theoretische Grundlagen

Mit der rasanten Verbreitung Neuer Medien in der Bildungslandschaft haben sich die Möglichkeiten der Präsentation von Lerninhalten vervielfältigt: Die Verquickung von Text, Bild, Ton und interaktiven Funktionen haben die Angebote belebt. Gleichzeitig stellt sich die Frage, ob und unter welchen Umständen der Einsatz dieser Möglichkeiten Vorteile beim Lernen bringt. Das interessiert insbesondere dort, wo ein virtuelles Lernangebot einen Inhaltsbereich abdecken soll, der eher fakten- denn prozessorientiert ist und deshalb herkömmlicherweise durch Printmedien abgedeckt wird. E-Learning hat durch seine Möglichkeiten, Interaktionen und Simulationen einzubauen, eindeutige Stärken, wenn es um die Darstellung und Schulung von Prozessen geht. Kann ein virtuelles Lernangebot aber auch eine Alternative für herkömmliche Textbücher darstellen und diese hinsichtlich der Leichtigkeit, mit der die Inhalte aufgenommen werden sogar übertreffen? Diese Frage lässt sich aus Sicht der Autoren unter der Bedingung bejahen, dass Texte und Abbildungen lernfreundlich und mit Rücksicht auf Stärken und Schwächen des Lernens am Computer gestaltet und das Verständnis dieser Inhalte mit den zur Verfügung stehenden Mitteln unterstützt wird. Sind diese Voraussetzungen erfüllt, lassen sich sowohl die Effizienz (ausgedrückt in benötigter Lernzeit bis zum Erreichen subjektiver Gewissheit, den Stoff zu beherrschen) als auch die Effektivität (ausgedrückt im Anteil korrekter Lösungen) gegenüber herkömmlichem Printmaterial steigern.

Für die vorliegende Studie wurden drei der in PTO verwendeten Darstellungsformen ausgewählt und einer empirischen Prüfung unterzogen:

- 1) eine interaktive Form der Text/Bild-Kombination (interaktive Grafik),
- 2) ein Pop-up-Glossar für Fachbegriffe und
- 3) eine Hinterleuchtungsfunktion zur Hervorhebung der relevantesten Textstellen auf einer Seite (Speed-Reading).

Im Folgenden wird zuerst eine theoretische Begründung für die Verwendung dieser Techniken präsentiert. Danach werden – getrennt für jede Technik – die in der Studie verwendeten Lernmaterialien, die Methoden und die Resultate erläutert. Die Befunde werden in einer Gesamtdiskussion kommentiert.

6.1.1 Bild/Text-Kombination (interaktive Grafik)

Hinsichtlich der Kombination verschiedener Repräsentationsformate legen Clark und Mayer (2000) eine kompakte und empirisch gut gestützte Sammlung von Empfehlungen vor. Aus ihren Leitlinien greifen wir für die vorliegende Studie eine Empfehlung heraus: Lerninhalte sollten möglichst in einer Kombination von Text und Bild dargestellt werden,

statt nur durch Text allein. Die Autoren nennen dies das "Multimedia-Prinzip". Der Ausdruck "Multimedia" ist allerdings nicht besonders scharf definiert: Mal wird darunter die Darbietung mittels verschiedener Geräte (Bildschirm, Lautsprecher) verstanden, mal multiple Codes und Modalitäten (Text, Standbilder, Film, Ton, haptische Reize usw.). Weidenmann (1994) empfiehlt denn auch eine Differenzierung des Begriffs in Multimodalität (im Sinne verschiedener Sinnesmodalitäten) und Multikodalität (im Sinne mehrerer Symbolsysteme). Da sich Text als auch (statische) Bilder sowohl am Computer wie auch in einem Buch darstellen lassen, rückt das Medium, in dem die Inhalte präsentiert werden, etwas in den Hintergrund. Es wird erst zentral, wenn Ton und/oder Bewegung eingesetzt werden sollen.

Für das Lernen deklarativer Inhalte scheint das Ansprechen mehrerer Sinneskanäle eine weniger grosse Rolle zu spielen als die Verwendung multipler Codes (Weidenmann, 1994). In einer Betrachtung der Effektivität und Effizienz des Lernens lohnt es sich deswegen, insbesondere auf den zweiten Aspekt zu fokussieren: die doppelte Kodierung durch Bild und Text. Im Zentrum des Interesses der vorliegenden Studie steht nicht die Multimedialität an sich, sondern der Vergleich zwischen einer monokodalen, rein textbasierten Darbietung und einer kombinierten Präsentation, die Text, Bildelemente und einfache Interaktionsmöglichkeiten verwendet und sich somit strukturell grundlegend von der rein textbasierten Darbietung unterscheidet.

Aus Sicht der Informationstheorie wäre nun zu erwarten, dass für einen Lerner, der mit doppelt kodierten Inhalten konfrontiert wird, auch ein Mehraufwand an Informationsverarbeitung und damit eine Verlängerung der Lernzeit resultiert. Aus der Perspektive der Kognitionspsychologie ist dies gerade nicht so. Hier geht man davon aus, dass eine simultane Nutzung sowohl der verbalen wie der bildlichen Verarbeitungskapazitäten eines Lerners den Lernprozess sogar beschleunigen kann. Im Folgenden werden die theoretischen Hintergründe für diese Annahme dargestellt und ein didaktisches Element einer Lernumgebung auf die postulierten Effekte hin überprüft.

Die hier dargestellte Untersuchung wurde im Zusammenhang mit Evaluationsmassnahmen der Online-Lernumgebung "Psychopathology Taught Online" der Universität Zürich, kurz "PTO", durchgeführt. PTO ist eine internetgestützte Lernumgebung mit dem Ziel, Lernenden im erweiterten Bereich der Psychologie die Symptomatik verschiedener psychischer Störungen zu vermitteln. Zum Zielpublikum gehören nicht nur Studierende der Psychologie, sondern auch Medizinstudierende, Ergotherapeutinnen, Psychiatriepfleger, Sozialpädagoginnen usw., Menschen also, die im Rahmen ihrer Berufstätigkeit regelmässig oder mit hoher Wahrscheinlichkeit mit psychischen Störungen konfrontiert werden.

Dieses Wissensgebiet wird in konventionellen Lehrbüchern stark textlastig vermittelt, da sich viele Symptome nur umständlich oder gar nicht visualisieren lassen (etwa „vermindertes Denk- oder Konzentrationsvermögen“ oder „Unschlüssigkeit als Symptome einer

depressiven Episode“). Trotzdem bestand bei der Konzeption von PTO das Ziel, die Stärken piktorialer Kodierung so oft wie möglich zu nutzen (wobei mit piktorialer Kodierung nicht nur realitätsnahe Abbilder, sondern auch logische Bilder im Sinne von Schnotz (1994) wie etwa Grafiken oder Diagramme gemeint sind). Dieses Bestreben beruht auf der Annahme, dass Menschen deklaratives Wissen vorwiegend über zwei zumindest zu Beginn der kognitiven Verarbeitung getrennte Kanäle aufnehmen können: den sprachlichen und den bildlichen (Mayer, 2001). Bereits Paivio (1986) geht davon aus, dass eine multiple Repräsentation einen Behaltensvorteil bringt. Text-Bild-Kombinationen sind zwar nicht prinzipiell und in jedem Fall einer monokodalen Repräsentation überlegen. Mittlerweile ist aber gut dokumentiert, unter welchen Bedingungen dies zutrifft (Mayer, 2005) (etwa, wenn die räumliche Nähe von Bildmaterial und korrespondierendem Text gewährleistet ist, Bild-Text-Scheren vermieden werden u. ä.). Entsprechende Prinzipien wurden bei der Gestaltung des im Folgenden noch beschriebenen Materials berücksichtigt.

Ein äusserst nützliches theoretisches Modell zur Erklärung der Prozesse beim Text- und Bildverstehen liefert Schnotz (2005). Es integriert verschiedene Modelle von Gedächtniskomponenten und kognitiven Verarbeitungsprozessen (Paivio, 1986; Baddeley, 1999). Dem Modell gemäss konstruiert ein Leser aufgrund des gelesenen Textes in seinem Arbeitsgedächtnis eine interne Repräsentation der Textbasis. Die darin enthaltenen Bedeutungen werden in Form von Propositionen gespeichert. Dies sind semantisch abgelegte Wissensbestände ohne Einkleidung in eine Alltagssyntax. Ein Bildbetrachter wiederum bildet eine analoge interne Repräsentation des Bildes und verdichtet dessen strukturelle Eigenschaften zu einem mentalen Modell mit bildhaftem Charakter. Letzteres bildet also die Entsprechung zu den Propositionen des verbalen Kanals. Diese beiden Repräsentationsformen – die Propositionen und die mentalen Modelle – stehen in Wechselbeziehung. Zum einen können gewisse Propositionen in Teile von mentalen Modellen "übersetzt" und diese dadurch ergänzt werden, zum anderen kann ein mentales Modell im Geiste "inspiziert" werden, woraus sich wiederum Propositionen bilden lassen. Dieses Modell vermag zu erklären, wieso Lernende einen besseren Lernerfolg erzielen mit Bild und Text statt nur mit Text allein – vorausgesetzt, die Kombination unterstützt eine widerspruchsfreie Bildung der beiden genannten Repräsentationen.

Geht man also davon aus, dass Lernende besonders dann einen guten Lernerfolg verzeichnen, wenn sie beides bilden können – einen möglichst umfassenden Satz von Propositionen und ein korrespondierendes mentales Modell – dann legt dies nahe, bei der Instruktion beides simultan anzubieten, sofern dies inhaltlich möglich und didaktisch gerechtfertigt ist. Diese Doppelkodierung sollte in einer wahrnehmungsgünstigen Form geschehen, die den Lernenden erlaubt, widerspruchsfreie Querverbindungen zu erkennen und die beabsichtigte Botschaft zu extrahieren.

6.1.2 Pop-up-Glossar

Einen Fliesstext zu lesen und zu verstehen, bedeutet wesentlich mehr als nur Worte zu erkennen und Konzepte im Gedächtnis zu aktivieren (Bock, 1978). Teilsätze müssen im Arbeitsgedächtnis zwischengespeichert und zueinander in Beziehung gesetzt werden. Die Strukturen, die aufgebaut werden, erfordern oft auch das Verknüpfen mehrerer Sätze. Das verbale Arbeitsgedächtnis wird dementsprechend stark belastet. Deshalb sollte dieser Prozess möglichst wenig unterbrochen werden. Andernfalls wird der Text als inhaltlich zusammenhangslos erlebt und man ist wiederholt gezwungen, ganze oder Teilsätze von neuem zu lesen (Ballstaedt, 1997). Solche Unterbrechungen können in Fachtexten oft durch komplexe Satzstrukturen entstehen, aber auch bei Verständnisproblemen, wenn sich die Notwendigkeit ergibt, ein unbekanntes oder nicht vertrautes Wort nachzuschlagen.

PTO verwendet in seinen Fliesstexten syntaktisch relativ einfache sprachliche Strukturen und bedient sich – verglichen mit manchen Lehrbüchern – eines möglichst unkomplizierten Vokabulars. In einer einführenden Lernumgebung sollen Fremdwörter aber nicht a priori vermieden werden. Vielmehr sollen die Lernenden ja insbesondere mit der Bedeutung der im Sachbereich geläufigen Fachtermini vertraut werden. Es stellt sich also das Problem, in welcher Form ihnen diese Fachbegriffe erklärt werden. Lehrbücher lösen dies oft durch ein Glossar am Ende des Buches, durch im Fliesstext platzierte Kästchen, manchmal auch mit Erläuterungen in einer speziell dafür vorgesehenen Spalte am Seitenrand. Im Printmedium bedeutet dies in der Regel, dass die Lesenden aus dem kontinuierlichen Lesefluss herausgerissen werden, insbesondere dann, wenn zum Nachschlagen eines Begriffs ein oder mehrere andere Bücher benutzt werden.

Diesem Umstand begegnet PTO mit der Verwendung einer Funktion der XML-basierten Sprache eLML (eLesson Markup Language, Fisler, 2007). Die Funktion – das Pop-up-Glossar – erlaubt es, Begriffe im Fliesstext farblich hervorzuheben. Fährt man mit der Maus über den Begriff, so wird in unmittelbarer Nähe eine Erläuterung eingeblendet. Das verhindert zum einen, dass die Lesenden Inhalte an anderer Stelle nachschlagen müssen und dabei gegebenenfalls mit anderen Inhalten ohne unmittelbaren Bezug zu ihrem Verständnisproblem konfrontiert werden. Zum andern bleiben sie auch örtlich auf die Textstelle fokussiert und können den Lesefluss sofort wieder aufnehmen, sobald die Begriffsbedeutung geklärt ist.

6.1.3 Textthinterleuchtung (Speed-Reading)

Wer viele Inhalte zu lernen hat, wird innert kürzester Zeit mit dem Problem konfrontiert, dass ihm oder ihr kaum genügend Zeit zur Verfügung steht, um alles zu sichten. Zudem müssen die Inhalte verdichtet oder auf eine andere Weise reduziert werden, weil man sich ohnehin kaum alles merken kann. Aus Textbüchern wird in der Regel meist nur aufge-

nommen, was einen zusammenhängenden Sinn stiftet und für die Lernziele relevant ist. Wenig relevanter Füllstoff wird sofort ausgeblendet, am Bildschirm noch deutlicher als in Büchern (Krug, 2006). Bei Texten als Präsentationsform – sei es im Buch oder am Bildschirm – spielt deshalb die Selektion von und die Reduktion auf als wichtig erkannte Aussagen eine zentrale Rolle. Gut lässt sich das erkennen, wenn man Lernende bittet, Zusammenfassungen von Texten zu schreiben. Selektion und Zusammenfassen sind grundlegende Reduktionsstrategien (Ballstaedt, 1997). In Printmedien wird dies in der Regel durch Zusammenfassungen am Ende eines Kapitels abgedeckt.

Dasselbe liesse sich auch in Bildschirmtexten umsetzen, mit dem gleichen Nachteil: Es ist zusätzlicher Text zu programmieren (respektive in Printmedien zu drucken). In PTO wurden stattdessen relevante Passagen, die zum Verständnis unerlässlich sind, mit speziellen Markierungen, sogenannten Tags, versehen. Diese erlauben es, auf Anklicken einer Schaltfläche hin, die entsprechend markierten Textstellen zu hinterleuchten und dadurch prominent hervorzuheben. Die Funktion ist in zweierlei Hinsicht nützlich: Wer zum ersten Mal eine Seite besucht, kann sich innert kürzester Zeit einen Überblick verschaffen, worum es im Text geht. Mit diesem "Vorwissen" lässt sich der Text dann wesentlich leichter im Detail lesen. Wer dagegen eine Seite zur Repetition besucht, klickt die Schaltfläche an und sieht sofort, was vom Autor/von der Autorin als wesentlich und zu memorieren erachtet wurde.

Allerdings erweist sich die Speed-Reading-Funktion nur dann als sinnvoll, wenn die hintereinander und ohne "Füllstoff" gelesenen, markierten Textstellen ein Minimum an Kohärenz ergeben. Beim Setzen der Anfangs- und Schluss-Tags der Markierungen wurde deshalb grosser Wert darauf gelegt, dass die Textstellen einigermaßen flüssig, semantisch kohärent und ohne syntaktische Widersprüche zu lesen sind. Im Sinne grösstmöglicher Sparsamkeit bei gleichzeitig kleinstmöglicher Informationseinbusse wurde dabei auf diverse Wörter verzichtet, die den Text normalerweise flüssiger lesbar machen. In der Praxis erweist es sich deshalb als einfacher, den ganzen Text zu lesen, es braucht aber auch deutlich mehr Zeit.

Welche Textstellen beim Einschalten der Speed-Reading-Funktion markiert werden, wurde sowohl nach inhaltlichen (Auswahl der in der Fachliteratur übereinstimmend betonten Sachverhalte) als auch nach sprachlichen Gesichtspunkten (Gewährleistung von semantischer und syntaktischer Kohärenz) bestimmt. Sie wurden so ausgewählt, dass lediglich durch das Lesen der gekennzeichneten Stellen der Sinn des Textes erfasst werden kann.

6.2 Experiment 1: Bild/Text-Kombination

6.2.1 Lernmaterial

Zur Untersuchung des didaktischen Elements interaktive Grafik wurde dessen Wirkung mit der Lernwirksamkeit einer Darstellung desselben Inhalts in einem konventionellen Buch verglichen. Dazu wurde eine interaktive Bild/Text-Kombination über die klassischen Angst- und Paniksymptome als Beispiel innerhalb der Lernumgebung PTO gewählt (Abb.44, links). Als Lehrtext in Buchform diente der entsprechende Ausschnitt aus der Internationalen Klassifikation der Krankheiten der WHO (ICD-10). Verwendet wurden die Forschungskriterien (Dilling, Mombour, Schulte-Markwort, Schmidt & 2005), weil sie den Inhalt in knapper, listenartiger Form abbilden, wie sie von Studierenden häufig für Lernzwecke benutzt wird. In den Kriterien mehrerer Angststörungen wird auf diese ICD-10-Liste der Angst- und Paniksymptome verwiesen. Darüber hinaus gehören phobische und andere Angststörungen zu den psychischen Störungen mit der vergleichsweise höchsten Verbreitung in der Bevölkerung. Die Liste kann deshalb als relevanter Lerninhalt angesehen werden. Es schien daher lohnenswert, diesen Satz von Symptomen medial besonders aufzubereiten. In der Lernumgebung PTO ist die Abbildung per Link in Form eines Popup-Fensters aufrufbar, jeweils auf der Seite der Kriterien der behandelten Angststörung (als Kriterien werden Angaben zu den zur Störung gehörenden Symptomen, ihrer Schwere und Dauer sowie zu den Ausschlusskriterien bezeichnet). Abb. 44 stellt die Darstellungsformate des Lerninhalts in den beiden Medien einander gegenüber.

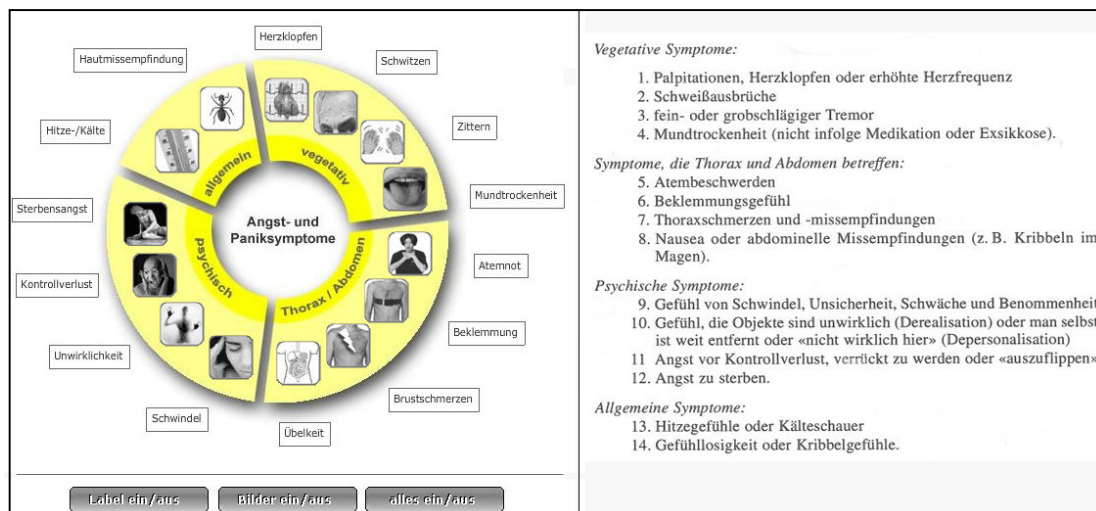


Abb. 44: Lernmaterial am Bildschirm (links) und im Buch (rechts).

Beim Vergleich der beiden Materialien wird der strukturelle Unterschied in der Präsentation deutlich: Die Text-Bild-Kombination hebt die sequenzielle Ordnung des Buchtextes

weitgehend auf. Die Symptome werden jeweils in einem stellvertretenden Begriff zusammengefasst, was der Lernstrategie des Reduzierens und Verdichtens entspricht (Ballstaedt, 1997). Die Symptomkategorien werden zur Benennung der Kreissegmente verwendet. Die Schaltflächen am Fuss der Grafik dienen dem stufenweisen Abdecken der Informationen: Zum Selbsttest können wahlweise alle Sammelbegriffe, die Sammelbegriffe samt Bilder und schliesslich sämtliche Inhalte einschliesslich der Kategoriebezeichnungen per Mausklick ausgeblendet werden.

Aus Abb. 44 ist erkennbar, dass in der konventionellen Textversion mehr Information zu stecken scheint als in der Text-Bild-Kombination der computergestützten Version. Diese Differenz wird dadurch aufgehoben, dass die Details in der Text-Bild-Kombination in einem Textfeld dargestellt werden, sobald die Lernenden auf die Bilder klicken (Abb. 45). Das Pop-up-Fenster der computergestützten Version enthält eine entsprechende kurze Bedienungsanweisung.

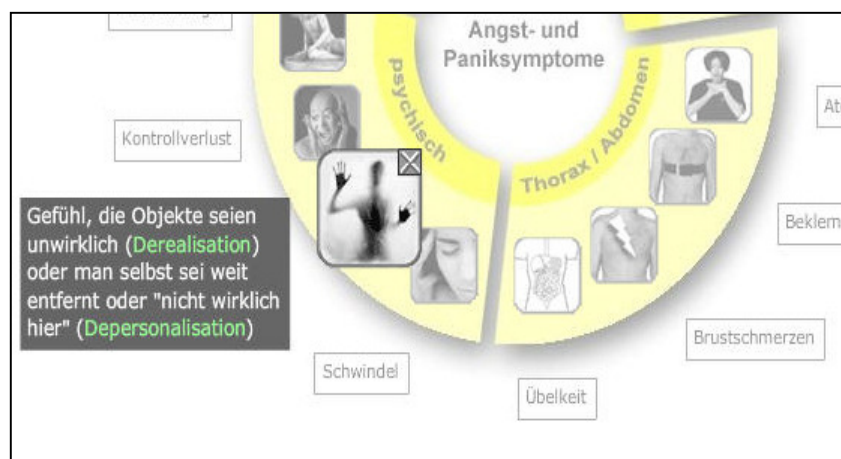


Abb. 45: Detailinformationen in der Text-Bild-Kombination.

Diese Lösung unterstützt das Lernen in dreierlei Hinsicht: Erstens wird die Information über die einzelnen Symptome in einem einzigen Begriff verdichtet. Zweitens wird der Inhalt an einen bildlichen Inhalt gekoppelt. Drittens wird damit eine erkundende Interaktion angeregt. Zusammen mit der oben erwähnten Ausblendefunktion zum Selbsttest bildet dies eine Handlungsmöglichkeit, die von Weidenmann (2001) neben dem Symbolsystem und der didaktischen Struktur als wesentliches Merkmal von Lernmedien hervorgehoben wird.

Im Beispiel der interaktiven Grafik wird der Lerninhalt auf zwei Ebenen doppelt kodiert: Zum einen haben die kleinen Abbilder eine unterstützende Funktion (doppelte Kodierung durch Sammelbegriff und symbolisierendes Bild). Zum andern vermittelt das segmentierte Kreismodell einen orientierenden Rahmen für die Symptombereiche und die Einteilung

der Symptome (doppelte Kodierung durch Kategoriebegriffe und geometrische Form). Beides zusammen bildet das mentale Modell, das in den Köpfen der Lernenden entstehen soll, ein Modell, an dem beim Abruf die Informationen gleichsam mental "abgelesen" werden können.

Der Wirksamkeitsvergleich einer Bild-Text-Kombination mit einer Nur-Text-Variante wäre natürlich auch in einem Printmedium denkbar. Mit den oben beschriebenen Möglichkeiten des Mediums Computer wird aber über diejenigen herkömmlicher Printmedien hinausgegangen: Erstens sind in der Abbildung selbst Interaktionen mit didaktischem Zweck möglich. Zweitens ist die Abbildung über Hypertext überall dort aufrufbar, wo aus den Kriterien einer Angststörung auf den entsprechenden Satz von Symptomen verwiesen wird.

6.2.2 Methode

Die oben eingeführte interaktive Grafik enthält die 14 massgeblichen Angst- und Paniksymptome in 4 Oberkategorien unterteilt. Untersucht werden sollte das Lernen mit der Angstkreis-Grafik im Vergleich zum herkömmlichen, buchbasierten Lernen. Das Lernziel im Versuch bestand im Aufzählen der Symptomklassen und der zugehörigen Symptome sowie in der korrekten Zuteilung der Symptome zu den Kategorien. Die Probanden sollten damit fähig sein, sämtliche gelernten Inhalte wiederzugeben. Je eine Lerngruppe lernte mit einem der beiden Formate. Eine Zeitvorgabe bestand nicht. Die Probanden wurden instruiert, die Inhalte so lange zu lernen, bis sie subjektiv der Überzeugung waren, diese vollständig wiedergeben zu können. Die Versuchspersonen wurden angehalten, vor Lernbeginn eine Stoppuhr zu starten und diese wieder zu stoppen, wenn sie sich sicher waren, den Inhalt fehlerfrei wiedergeben zu können. Das Format, in welchem die Probanden das Gelernte wiedergaben, sollte nicht zu dicht an einem der Lernformate liegen. Daher wurden die Probanden im Anschluss ans Lernen gebeten, die 4 Symptomklassen und die zugehörigen 14 Symptome auf Band zu sprechen. Die Audioaufnahmen wurden im Nachhinein vom Testleiter abgehört und auf Vollständigkeit und Korrektheit geprüft.

Die interaktive Grafik als didaktisches Element wurde bezüglich Lerneffizienz (gemessen über die Lernzeit) und Lernerfolg (gemessen über die Korrektheit der Resultate) untersucht. Dabei wurde mit folgenden Effekten gerechnet:

- 1) Die Lernenden gelangen mit Hilfe der doppelten Kodierung in PTO signifikant schneller zum subjektiven Gefühl, die Inhalte zu beherrschen und vollständig wiedergeben zu können (grössere Effizienz mit PTO).

- 2) Hinsichtlich der Fehlerquote bei der Wiedergabe wurden keine signifikanten Unterschiede erwartet, da das Lernmaterial auch in reiner Textform eine 100%ige Trefferquote erlaubte.

Für die Untersuchung wurden insgesamt 24 Personen getestet. Experimental- und Kontrollgruppe bestanden aus je 12 Personen. Als Probanden und Probandinnen wurden Personen aus der potenziellen Nutzergruppe von PTO gewählt. Studierende der Psychopathologie oder sonstige Personen, welche bereits über Vorwissen über die geprüften Inhalte verfügten, wurden nicht berücksichtigt.

6.2.3 Resultate

Abb. 46 zeigt die Lerndauer für die Experimentalgruppe, welche anhand der interaktiven Grafik gelernt hatte, sowie der Kontrollgruppe, welche sich die Lerninhalte mittels Lehrbuch angeeignet hatte. Die Experimentalgruppe benötigte durchschnittlich 361 Sekunden (Stdv: 109 s), die Kontrollgruppe dagegen 658 Sekunden (Stdv: 377 s), für das Lernen der Inhalte. Der Unterschied wurde auf dem 5%-Niveau signifikant ($t(13) = -2.62, p < .05$).

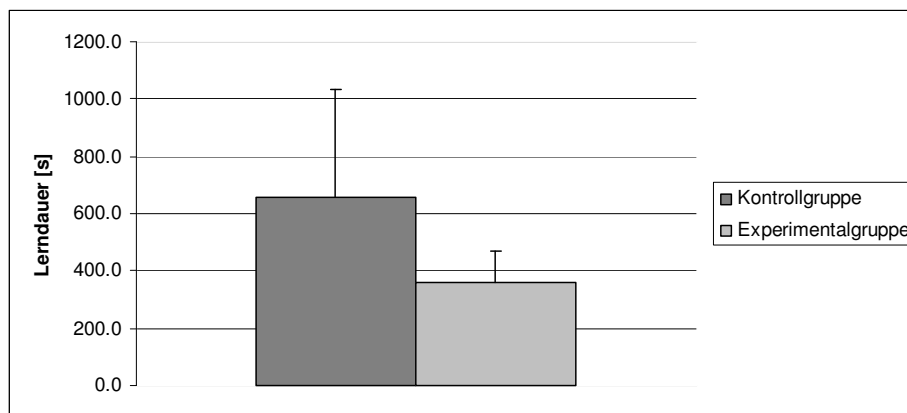


Abb. 46: Durchschnittliche Lerndauer der Kontrollgruppe (gelernt mit Lehrbuch) sowie der Experimentalgruppe (gelernt mit interaktiver Grafik).

Abb. 47 zeigt die durchschnittliche Anzahl korrekter Antworten für die Experimental- und die Kontrollgruppe. Die Gesamtzahl möglicher korrekter Antworten lag bei 18. Die Experimentalgruppe erzielte durchschnittlich 16.3 korrekte Antworten (Stdv: 2), während deren Zahl bei der Kontrollgruppe 15.8 (Stdv: 1.6) betrug. Der gefundene Unterschied zwischen den beiden Gruppen ist nicht signifikant ($t(22) = 0.56, p > .05$).

Bezüglich der Anzahl nicht korrekter Antworten (falsche und nicht gewusste Antworten) konnte zwischen den Gruppen kein signifikanter Unterschied festgestellt werden.

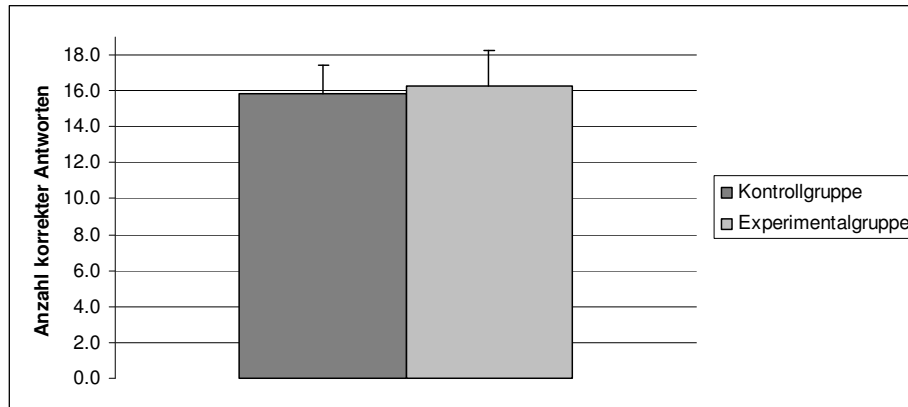


Abb. 47: Durchschnittliche Anzahl korrekter Antworten der Kontrollgruppe (gelernt mit Lehrbuch) sowie der Experimentalgruppe (gelernt mit interaktiver Grafik).

6.3 Experiment 2: Pop-up-Glossar

6.3.1 Lernmaterial

Für die Untersuchung der Wirkung des Pop-up-Glossars wurde innerhalb von PTO ein Textabschnitt über grundlegende Merkmale der Schizophrenie ausgewählt. Der Text behandelte insbesondere Möglichkeiten zur Kategorisierung der Symptome sowie prototypische Erkrankungsstadien. Er hatte eine Gesamtlänge von ca. 2.5 Bildschirmseiten und enthielt insgesamt 30 fachspezifische Termini, die mit dem interaktiven Glossar versehen waren. Wird der Mauszeiger über einen solchen Fachbegriff platziert, erscheint die Definition in einem kleinen Pop-up-Fenster unmittelbar neben dem fraglichen Begriff angezeigt. Die Begriffe, für die eine solche Definition zur Verfügung steht, sind zur leichten Erkennbarkeit im Text grün markiert. Dadurch lässt sich bei Bedarf kontextspezifisch auf Informationen zurückgreifen, wodurch sich Unterbrechungen und Störungen im Leseprozess vermeiden lassen und dieser insgesamt effizienter wird (siehe Abb. 48).

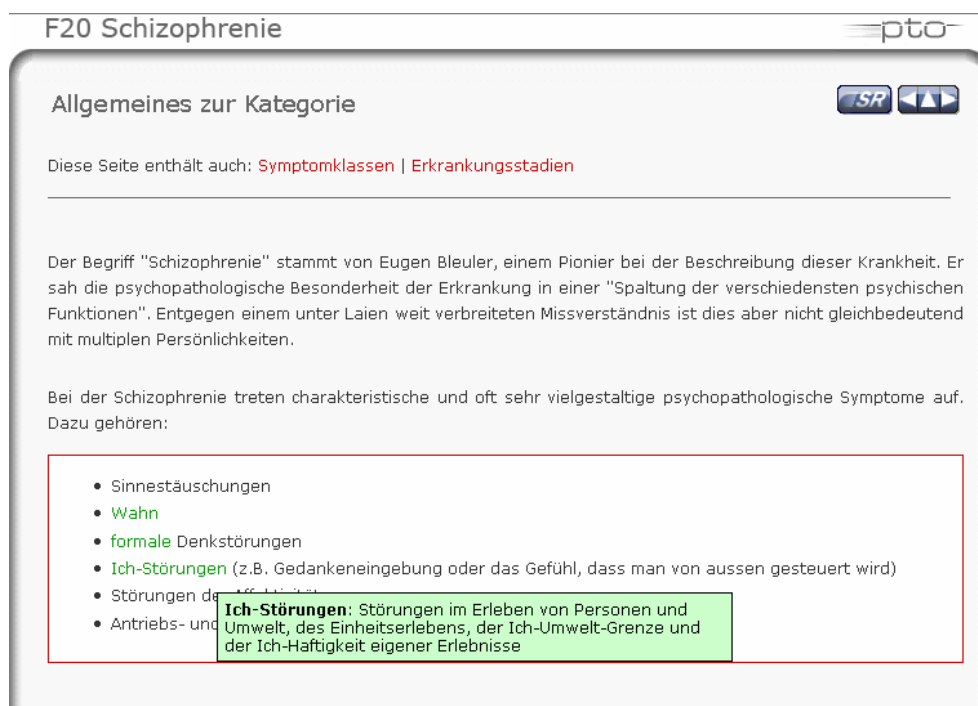


Abb. 48: Textausschnitt aus PTO zur Schizophrenie mit eingeblendeter Definition des Begriffs „Ich-Störungen“.

Der Text strukturierte sich in einen ersten Teil mit der Darstellung der relevantesten Symptome sowie Erläuterungen zu den Symptomklassen und Symptomkategorisierungen. Der zweite Textteil enthielt Informationen zu den verschiedenen Krankheitsstadien.

6.3.2 Methode

Die Probanden wurden instruiert, sich mit der Seite mit den allgemeinen Informationen zur Schizophrenie auseinanderzusetzen. Sie sollten den Text lesen und sich dabei mit den Definitionen der grün markierten Begriffe soweit vertraut machen, dass sie diese nachträglich erklären konnten. Die eine Experimentalgruppe konnte dabei direkt auf die im Pop-up-Glossar präsentierten Definitionen zurückgreifen, während die andere Gruppe verschiedene Lexika und Kompendien zur Verfügung gestellt bekam, in denen die entsprechenden Begriffe nachgeschlagen werden konnten. Die Probanden wurden angehalten, nicht bloss die Definitionen zu lernen, sondern sich auch mit den restlichen Inhalten der Seite zu beschäftigen und damit den Kontext, in dem die definierten Begriffe standen, mitzuerfassen. Das Lernziel bestand darin, dass die Probanden die auf der Seite enthaltenen Definitionen kannten und erklären konnten. Darüber hinaus sollten sie erläutern können, in welchem Kontext der Begriff genannt wurde. Eine Zeitvorgabe bestand grundsätzlich nicht, jedoch wurde ein oberes Limit von 30 Minuten festgelegt und vor Testbeginn kommuniziert. Geprüft wurden schliesslich fünf Begriffsdefinitionen, und zwar mit einer Methode, welche sich an der Heidelberger Strukturlegetechnik orientierte. Dabei wurden den Probanden die Kernbegriffe aus den einzelnen Definitionen auf Kärtchen gedruckt vorgelegt. Insgesamt standen 25 Kärtchen zur Verfügung, auf denen jeweils ein bis zwei Worte standen, die eine semantische Einheit bildeten (z. B. "flexibilitas cerea" oder "nicht korrigierbar"). Die Aufgabe der Probanden bestand darin, die einzelnen Kärtchen zu sinnvollen Definitionen zusammenzufügen. Dazu stand ihnen ein grosses Blatt mit fünf aufgezeichneten ovalen Bereichen zur Verfügung, die mit den gefragten Fachbegriffen bezeichnet waren. In diese Bereiche sollten zwischen vier und fünf Kärtchen so platziert werden, dass sich unter Verwendung der in einem Bereich befindlichen Begriffe eine stimmige Definition des Fachbegriffs bilden liess, zu dem man die Kärtchen gelegt hatte. Es ging also darum, die Kärtchen sinnvoll auf die verschiedenen Fachbegriffe zu verteilen.

Indem die Probanden dem Testleiter die Definitionen anhand der gelegten Kärtchen ausformulierten, fand eine Form der kommunikativen Validierung statt. Gemessen wurde zum einen die Zeit, bis die Probanden erklärten, dass sie die Begriffsdefinitionen beherrschten (Lerneffizienz). Zum andern wurde die Anzahl korrekt verwendeter Elemente gezählt (Lerneffektivität). Die gestellte Hypothese ging von zweierlei Unterschieden im Vergleich zur Darstellung desselben Inhalts mittels Printmedien aus:

- 1) Die Lernenden entwickeln unter Verwendung des Pop-up-Glossars im Mittel signifikant schneller das subjektive Gefühl, den angestrebten Stoff zu beherrschen (grössere Effizienz mit PTO).
- 2) Den Lernenden gelingen in einer Verständnisaufgabe signifikant mehr korrekte Begriffskategorisierungen (grössere Effektivität).

Getestet wurden dieselben Probanden wie bei den Erhebungen zur interaktiven Grafik. Die Reihenfolge, in dem die beiden Tests durchgeführt wurden, war über die gesamte Stichprobe ausbalanciert.

6.3.3 Resultate

Abb. 49 zeigt die Lerndauer für die Experimentalgruppe, welche anhand der computergestützten Version gelernt hatte, sowie der Kontrollgruppe, welche die Definitionen in den Lehrbüchern und Kompendien nachgeschlagen hatte. Die Experimentalgruppe benötigte durchschnittlich 1152 Sekunden (Stdv: 420 s), die Kontrollgruppe dagegen 1557 Sekunden (Stdv: 279 s) für das Lernen der Definitionen. Der Unterschied ist auf dem 5%-Niveau signifikant ($t(22) = -2.78, p < .01$).

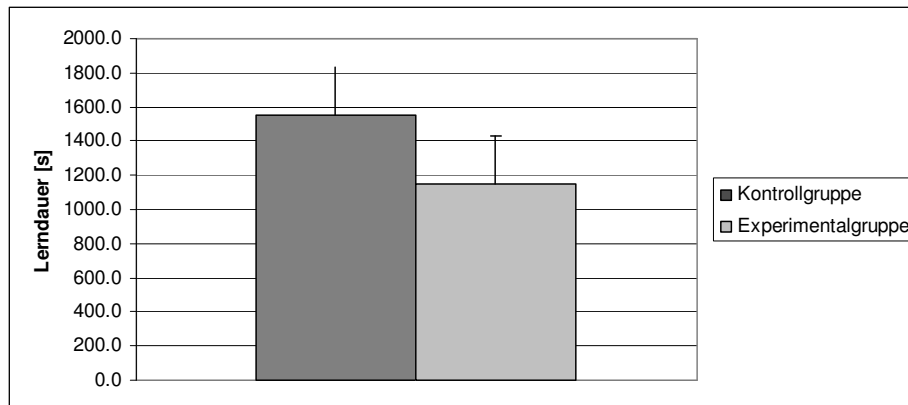


Abb. 49: Durchschnittliche Lerndauer der Kontrollgruppe (gelernt mit Lehrbuch) sowie der Experimentalgruppe (gelernt mit Pop-up-Glossar).

Abb. 50 zeigt die durchschnittliche Anzahl korrekter Antworten für die Experimental- und die Kontrollgruppe. Die Gesamtzahl möglicher korrekter Antworten lag bei 23. Die Experimentalgruppe erzielte durchschnittlich 16.4 korrekte Antworten (Stdv: 3.6), während deren Zahl bei der Kontrollgruppe 11.3 (Stdv: 3.5) betrug. Der gefundene Unterschied zwischen den beiden Gruppen ist ebenfalls signifikant ($t(22) = -3.53, p < .01$).

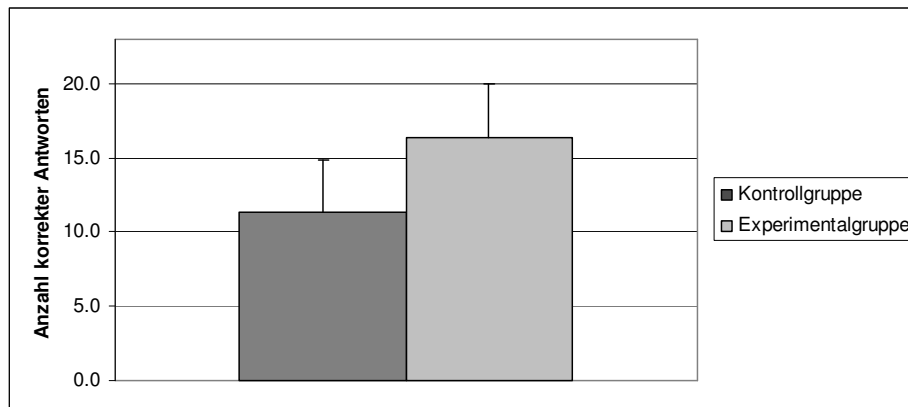


Abb. 50: Durchschnittliche Anzahl korrekter Antworten der Kontrollgruppe (gelernt mit Lehrbuch) sowie der Experimentalgruppe (gelernt mit Popup-Glossar).

6.4 Experiment 3: Speed-Reading

6.4.1 Lernmaterial

Wie oben bereits erläutert handelt es sich beim Speed-Reading um eine spezifisch im Kontext von PTO entwickelte Funktion. Analog zum Papierlernen, bei dem die Lernenden häufig mittels Markierstiften die besonders relevanten Stellen hervorheben, bietet PTO die Möglichkeit für eine digitale Kennzeichnung. Hierzu ist in PTO über den entsprechenden Texten ein Button platziert. Wird dieser aktiviert, erscheinen die zentralen Textstellen leuchtend gelb hinterlegt. Dies soll es den Lernenden ermöglichen, ihren Lernprozess effizienter und effektiver zu gestalten.

Zwei Überlegungen haben zur Implementierung dieser Funktion geführt: Die Lernenden sollen die Gelegenheit haben, beim Erstkontakt mit PTO-Texten auf Wunsch vorab einen Grobüberblick über die jeweiligen Inhalte zu gewinnen, bevor sie den Text im Detail studieren (ähnlich einem Advance Organizer im Sinne Ausubels). Zum andern soll das Wiederholen erleichtert werden, indem das Wesentliche bei einem späteren Besuch der Seite noch einmal hervorgehoben wird. Die Verdichtung, die man in einem Lehrbuch oder in einer gedruckten Version mittels Textmarker erreicht, soll bereits im elektronischen Format am Bildschirm verfügbar sein. Allerdings besteht nicht die Absicht, die Lernenden nur die markierten Stellen lesen zu lassen. Markierte Stellen wurden so gewählt, dass die wesentliche Aussage des Textes auf unspezifischer Ebene zusammengefasst wird. Nicht markierte Stellen dagegen enthalten Informationen, welche die markierten Stellen akzentuieren, ihre Bedeutung präzisieren, in einen Zusammenhang stellen oder Beispiele dazu liefern. Darüber hinaus liefern sie den "Leim", der die markierten Stellen zusammenhält und dem Text einen Lesefluss verleiht. Solche Stellen sind in dieser Textsorte deshalb keineswegs überflüssig (anders als etwa bei Listen oder tabellarischen Darstellungen, wo die Kürze Priorität hat).

Für die Studie wurde ein Text zum Thema Borderline-Störung gewählt (Abb. 51 zeigt denselben Textausschnitt ein Mal mit sowie ein Mal ohne aktivierte Speed-Reading-Funktion).

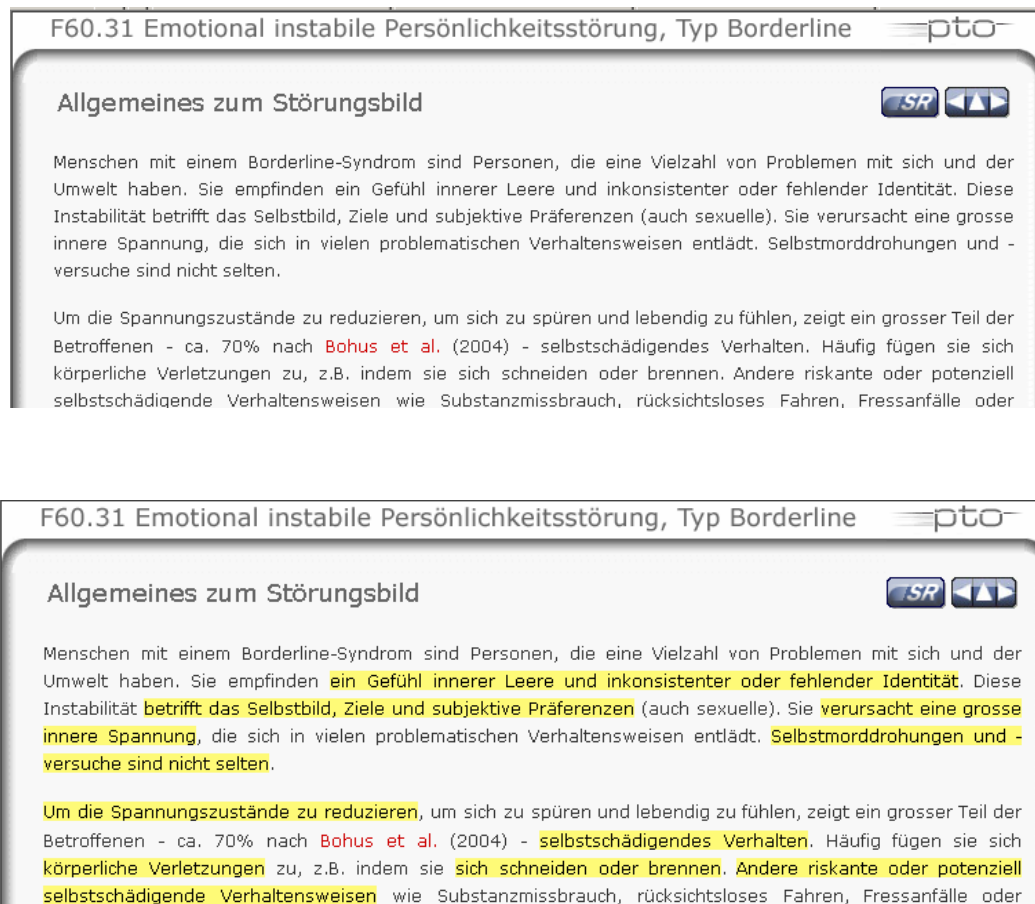


Abb. 51: Text zur Borderline-Störung ohne und mit aktiviertem Speed-Reading.

Der Text beansprucht ca. 1.5 Bildschirmseiten. Im Sinne einer Einführung beinhaltet er allgemeine Informationen zur emotional instabilen Persönlichkeitsstörung des Typs Borderline. Erläutert werden darin die wesentlichen Merkmale der Störung.

6.4.2 Methode

In einer Vorstudie wurde ermittelt, wie viel Zeit das bloße Durchlesen des Texts zur Borderline-Störung in der Volltextversion im Mittel beansprucht (360 Sekunden). Danach wurde der prozentuale Anteil des Speed-Reading-Textes am Volltext bestimmt. Derselbe Prozentsatz wurde nun von der Zeit für das Durchlesen genommen (170 Sekunden).

Als Testmaterial diente der oben erwähnte Text zur Borderline-Störung, welcher den Probanden zum Durchlesen vorgelegt wurde. Getestet wurde unter 3 Bedingungen: Die erste Bedingung beinhaltete das Lesen des Textes in 170 Sekunden mit aktivierter Speed-Reading-Funktion. Dabei sollten sich die Probanden nur auf die gekennzeichneten Textstellen konzentrieren. Gemäss der Vorstudie sollten diese 170 Sekunden gerade für das einmalige Durchlesen der markierten Textstellen reichen.

Die zweite Bedingung bestand im Lesen des Volltextes ohne aktivierte Speed-Reading-Funktion in 360 Sekunden. Die dritte Bedingung schliesslich umfasste das Lesen des Textes in 360 Sekunden mit aktivierter Speed-Reading-Funktion. Bei dieser Bedingung wurden die Probanden angehalten, während der ersten 240 Sekunden zügig den Volltext zu lesen, um dann während der verbleibenden 120 Sekunden nur auf die gekennzeichneten Stellen zu fokussieren.

Im Anschluss an das Lesen wurde das Textverständnis der Probanden anhand eines Lückentexts geprüft. Dieser bestand aus 3 kurzen Absätzen, enthielt insgesamt 24 Lücken und nahm inhaltlich Bezug auf die im Speed-Reading gekennzeichneten Stellen. Bei der Zusammenstellung des Lückentextes wurde darauf geachtet, dass die Anforderungen selbst für mit der Materie vertraute Personen hoch waren. Dadurch sollte ein möglicher Deckeneffekt verhindert werden. Weiter sollte auf Grund der Lektüre des Textes eindeutig klar sein, welches Wort oder welche Wörter bei den jeweiligen Lücken gesucht waren. Die Probanden waren frei in ihrer Wortwahl, Bedingung war aber die inhaltliche Korrektheit und syntaktische Kohärenz. Die Beurteilung der inhaltlichen Korrektheit erfolgte durch einen unabhängigen Fachmann. Dieser bekam sämtliche Antworten vorgelegt und beurteilte diese analog zu Prüfungsergebnissen. Er befand, ob die eingesetzten Begriffe semantisch zulässig waren und damit als korrekt gewertet werden konnten. Darüber hinaus wurden drei Lücken, die sich im Testverlauf als zu wenig eindeutig erwiesen, nachträglich aus der Auswertung ausgeschlossen.

Das Experiment verfolgte zwei Ziele: Zum einen sollte untersucht werden, wie sich das Speed-Reading auf die Lerneffizienz auswirkt. Hierzu diente der Vergleich der Bedingungen 1 und 2 (Speed-Reading in 170 Sekunden vs. Volltext in 360 Sekunden). Zum anderen galt es zu untersuchen, ob sich der Lernerfolg bei gleicher Zeitvorgabe durch Speed-Reading erhöhen lässt. Verglichen wurden dabei die Bedingungen 2 und 3 (Volltext in 360 Sekunden vs. Speed-Reading in 360 Sekunden). Als kritisches Mass diente die Anzahl korrekter Antworten. Folgende Hypothesen wurden im Zusammenhang mit dem Einsatz der Speed-Reading-Funktion gestellt:

- 1) Die Versuchspersonen, die den Text mit Speed-Reading in 170 Sekunden lesen (Bedingung 1), schneiden im Test gleich gut ab wie die Leser der Volltextversion (Bedingung 2). Bezüglich der Testleistung wurden somit keine signifikanten Unterschiede erwartet (grössere Effizienz).

- 2) Die Probanden, welche den Text mit aktivierter Speed-Reading-Funktion in 360 Sekunden lesen (Bedingung 3), erzielten signifikant bessere Resultate als die Leser der Volltextversion ohne Speed-Reading-Markierung (Bedingung 2) (grössere Effektivität).

Getestet wurden wiederum dieselben Probanden wie bei der Untersuchung der interaktiven Grafik und des Pop-up-Glossars. Die Untersuchung des Speed-Readings enthielt jedoch eine dritte Bedingung und verlangte damit nach einer zusätzlichen Experimentalgruppe von 12 Personen. Insgesamt wurden damit 36 Personen getestet. Für jene Probandengruppe, welche sämtliche Experimente durchlief, wurde die Reihenfolge, in der die didaktischen Elemente bearbeitet wurden, ausbalanciert.

6.4.3 Resultate

Abb. 52 zeigt die durchschnittliche Anzahl korrekter Antworten für alle drei Bedingungen. Die maximal erreichbare Punktzahl lag bei 21. Die Probanden der Bedingung 1 (Speed-Reading; 170 Sekunden) erzielten im Mittel 9.2 korrekte Antworten (Stdv: 3.4), während diejenigen der Bedingung 2 (Volltext; 360 Sekunden) 11.8 (Stdv: 3.0) erreichten. Die Probanden der Bedingung 3 erzielten im Durchschnitt 15 korrekte Antworten (Stdv: 3.8).

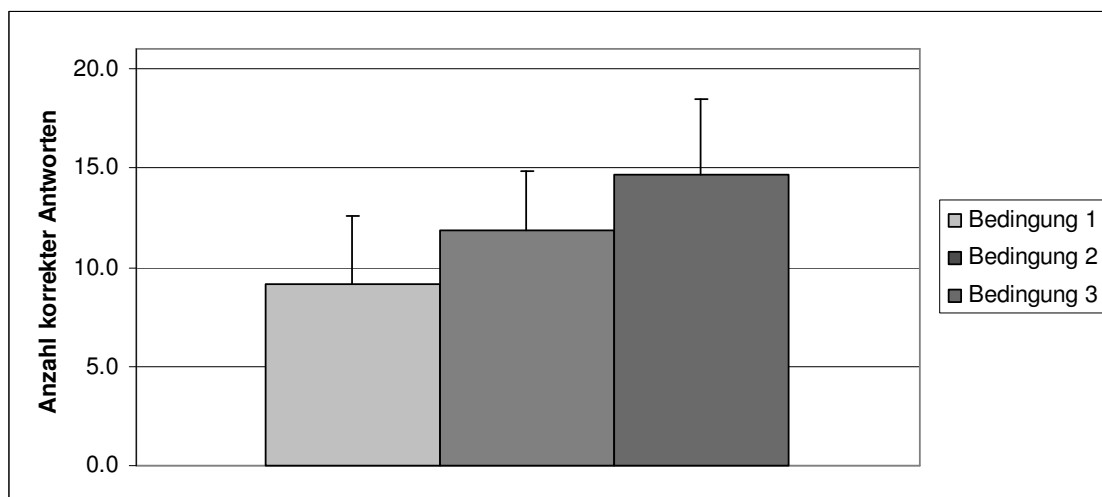


Abb. 52: Durchschnittliche Anzahl korrekter Antworten unter der Bedingung 1 (Speed-Reading; 170 Sekunden), Bedingung 2 (Volltext; 360 Sekunden) sowie der Bedingung 3 (Speed-Reading; 360 Sekunden).

Die statistische Überprüfung mittels Varianzanalyse (ANOVA) ergab einen signifikanten Unterschied auf Gruppenebene, $F(2, 33) = 7.88$, $MSE = 90.78$, $p < .01$.

Für die Einzelvergleiche zeigte sich eine signifikante Differenz zwischen den Bedingungen 1 (Speed-Reading; 170 Sekunden) und 3 (Speed-Reading; 360 Sekunden). Diese wurde auf dem 1%-Niveau signifikant. Die beiden anderen Unterschiede zwischen den Bedingungen 1 und 2 sowie zwischen den Bedingungen 2 und 3 erwiesen sich als nicht signifikant (siehe Tab. 8).

Tab. 8: Resultate der Post-Hoc-Tests zum Vergleich der Mittelwerte.

| | Bedingung 1 | Bedingung 2 | Bedingung 3 |
|-------------|--------------------------------|--------------------------------|-------------|
| Bedingung 1 | | | |
| Bedingung 2 | Nicht signifikant $p > .05$ | | |
| Bedingung 3 | Signifikant $p < .01$ | Nicht signifikant $p > .05$ | |

6.5 Diskussion

Im Fokus der Untersuchung stand die Frage nach Lerneffizienz und Lernerfolg. Postuliert wurde im Allgemeinen, dass die Experimentalgruppe (Lernen mit den Funktionen von PTO) im Unterschied zur Kontrollgruppe (Lernen mit Lehrbuch oder unmodifiziertem Text) effizienter und effektiver lernt. Als Effizienzsteigerung gilt dabei ein Zeitgewinn bei mindestens gleicher Leistung im Test, als Effektivitätssteigerung ein besseres Testergebnis. Zwei Versuche – der eine zur interaktiven Grafik und der andere zum Pop-up-Glossar - waren als einfache Vergleiche zwischen Experimental- und Kontrollbedingung angelegt. Sie sollen im Folgenden zuerst kommentiert werden. Danach folgen Erläuterungen zur Untersuchung der Speed-Reading-Funktion, bei der drei Bedingungen miteinander verglichen wurden.

Interaktive Grafik. Beim Versuch mit der grafischen Darstellung der Paniksymptome stand vor allem die Frage nach der Lerneffizienz im Zentrum. Diejenigen Probanden, welche mit der interaktiven Grafik lernten, waren im Vergleich zu den Lehrbuch-Lernern in der Tat fast doppelt so schnell und schnitten im anschliessenden Test - wenn auch nicht signifikant - leicht besser ab. Dass bei der Reproduktion (Lernerfolg) kaum Unterschiede in der Trefferzahl zu verzeichnen waren, kann daran liegen, dass das Lernmaterial auch in der konventionellen Textversion zu geringe Anforderungen stellte, als dass die Vorteile der Doppelkodierung ihre Wirkung auch hinsichtlich der Effektivität hätten entfalten können. Bei komplexerem und umfangreicherem Lernmaterial wären hier allenfalls deutlichere Unterschiede zu erwarten gewesen. Diese Lesart wird durch den gefundenen Deckeneffekt untermauert.

Unsere Resultate weisen darauf hin, dass die doppelte Kodierung des Lernmaterials in Bild- und Textform der Lerneffizienz förderlich ist. Diese doppelte Kodierung wird für den Produzenten des Lernmaterials zwar zu erhöhtem Aufwand führen, die Rezipienten aber profitieren davon, dass ihnen die Bildung eines mentalen Modells durch die zusätzliche Information erleichtert wird. Vor einem informationstheoretischen Hintergrund mag dies erstaunen, ausgehend von einem kognitionspsychologischen Modell einer doppelten Abspeicherung ist dies aber durchaus verständlich.

Pop-up-Glossar. Die empirischen Daten zur Untersuchung im Zusammenhang mit dem Pop-up-Glossar bestätigen beide getätigten Hypothesen. Die Probanden, welche die Definitionen der Fachbegriffe mittels elektronischem Medium lernten, erzielten die deutlich bessere Testleistung als die Lehrbuch-Lerner und benötigten hierfür zudem bedeutend weniger Zeit. Dies mag vor allem daran liegen, dass das Lernen mit dem Pop-up-Glossar zu keinen Störungen im Lesefluss führte. Die Lehrbuch-Lerner waren durch das ständige Nachschlagen in verschiedenen Lexika und Kompendien zu Unterbrechungen im Lern-

prozess gezwungen, was sich negativ auf die Lerneffizienz als auch insgesamt nachteilig auf die Gedächtnisleistung ausgewirkt zu haben scheint.

Das Glossar-basierte Lernen bietet zudem die Möglichkeit, innert Kürze auf eine bereits gelernte Begriffsdefinition zurückzugreifen, während dies beim Lehrbuchlernen nur mit grösserem Aufwand möglich ist. Die Repetition einzelner Definitionen fällt dadurch um einiges einfacher. Weiter kann sich die konsistente Art und Weise der Aufbereitung der Definitionen (im Stil der Darstellung sowie der Art der prägnanten Ausformulierung) in PTO im Gegensatz zu der grossen Spannbreite der Umsetzung in den verschiedenen Lehrbüchern vorteilhaft auf die Lernleistung ausgewirkt haben.

Speed-Reading. Die erste getestete Bedingung diente dem Vergleich der Textbearbeitung mit Speed-Reading-Funktion mit dem Lesen der Volltextversion. Den Probanden, welche die Speed-Reading-Funktion verwenden konnten, stand nicht genügend Zeit zur Verfügung, um den gesamten Text zu lesen. Sie mussten sich daher rein auf die markierten Textstellen konzentrieren. Geprüft wurde hierbei bezüglich der Testleistung (Lerneffektivität). Die Testresultate im Lückentext fielen zwar nicht signifikant, aber tendenziell etwas schlechter aus, was zwar nicht ganz im Sinne der Hypothese war (sie postulierte gleiche Leistungen im Test), aber trotzdem plausibel scheint. Der Unterschied mag auf die mehr als doppelte Zeit zurückzuführen sein, mit der sich die Lernenden in der Volltextbedingung ohne Markierungen mit dem Text beschäftigen konnten. Vermutlich spielt hier aber auch der Inhalt der nicht markierten Textstellen eine Rolle, der dem Text mehr Kohärenz verlieh und die Chance erhöhte, dass sich die Probanden das Beschriebene besser merken konnten.

Die zweite geprüfte Bedingung umfasste den Vergleich zwischen dem Lesen der Volltextversion ohne Speed-Reading mit der zumindest teilweisen Nutzung von Speed-Reading (240 Sekunden Lektüre der Volltextversion mit anschliessender erneuter Bearbeitung mit Speed-Reading während 120 Sekunden). Hierbei bestand die Erwartung, dass die Betonung der relevanten Stellen, auf die sich konsequenterweise auch der Lückentext bezog, einen grösseren Lernerfolg zeigen würde. Allerdings wurde das 5%-Signifikanzniveau (mit $p = .07$) knapp verpasst. Es darf erwartet werden, dass eine mit einer grösseren Stichprobe angelegte Untersuchung zu einem signifikanten Unterschied zwischen diesen beiden Bedingungen geführt hätte.

Dass das Miterfassen der nicht markierten Stellen nicht zu vernachlässigen ist, zeigt der signifikante Unterschied zwischen der Bedingung mit ausschliesslichem Speed-Reading und der Bedingung mit teilweiser Nutzung von Speed-Reading (bei mehr zur Verfügung gestellter Zeit). Die Hauptunterschiede zwischen den beiden Bedingungen lagen hier in der vorgeschalteten Volltext-Lektüre und in der wiederholten Lektüre der markierten, relevanten Textteile. Aufgrund des Untersuchungsdesigns lässt sich nicht genau sagen,

welcher der beiden Faktoren welchen Einfluss ausübt. Eine Nachfolgeuntersuchung würde sich zur Klärung dieser Frage anbieten.

Die Resultate zeigen, dass sich mit Speed-Reading allein bereits ein gewisser Lernerfolg erzielen lässt. Dies spricht für einen Effizienzgewinn durch Speed-Reading, wie es hier implementiert wurde. Allerdings war diese artifizielle Bedingung (ausschliessliches Speed-Reading) nie für das "echte" Lernen mit PTO vorgesehen, oder doch höchstens als Vorbereitung für die Volltext-Lektüre. Die relevantere Hypothese bezog sich auf eine Steigerung der Effektivität – höherer Lernerfolg bei gleicher Zeitinvestition – durch die Addition von Speed-Reading. Dieser Effekt konnte nicht in der gewünschten Deutlichkeit demonstriert werden, Tendenzen sind aber erkennbar.

Insgesamt haben die Untersuchungen erwiesen, dass die in PTO eingesetzten didaktischen Elemente einen positiven Effekt auf den Lernprozess ausüben. Dieser beruht in erster Linie in einer Effizienzsteigerung, umfasst aber auch eine Zunahme der Effektivität. Der Gewinn an Effizienz und Effektivität wurde dabei durch die neuen Darstellungsformen und technischen Möglichkeiten erzielt, welche sich mit dem Einsatz eines elektronischen Lernmediums ergaben. Richtig eingesetzt führen diese zu einer verdichteten und optimierten Aufbereitung und Präsentation der Inhalte, welche es den Lernenden erlaubt, schnell und ohne Unterbrechung im Lernprozess auf die im jeweiligen Lernkontext relevanten Informationen zuzugreifen. E-Learning ermöglicht damit den Lernenden im Vergleich mit herkömmlichen Lehrbüchern des Fachgebiets eine insgesamt effizientere und effektivere Gestaltung des Lernvorgangs.

6.6 Literaturverzeichnis

- Baddeley, A. D. (1999). *Essentials of Human Memory*. Hove: Psychology Press.
- Ballstaedt, S. P. (1997). *Wissensvermittlung*. Weinheim: Beltz Psychologische Verlagsunion.
- Bock, M. (1978). *Wort-, Satz-, Textverarbeitung*. Stuttgart: Kohlhammer.
- Clark, R. & Mayer, R. E. (2000). *e-Learning and the Science of Instruction*. San Francisco, Pfeiffer.
- Dilling, H., Mombour, W., Schmidt, M. H. & Schulte-Markwort, E. (2005). *Internationale Klassifikation psychischer Störungen. Diagnostische Kriterien für Forschung und Praxis*. Bern: Verlag Hans Huber.
- Fisler, J. (2007). Erstellung von strukturierten e-Learning Inhalten mit eLML (eLesson Markup Language). ZInfo - Die elektronische Zeitschrift der Informatikdienste [On-line]. Available: <http://www.id.uzh.ch/publikationen/zinfo/>
- Krug, S. (2006). *Web-Usability. Das intuitive Web*. Heidelberg: Redline.
- Mayer, R. E. (2001). *Multimedia Learning*. Cambridge: Cambridge University Press.
- Mayer, R.E. (2005). *The Cambridge Handbook of Multimedia Learning*. Cambridge: Cambridge University Press.
- Paivio, A. (1986). *Mental Representations. A Dual Coding Approach*. New York: Oxford University Press.
- Schnotz, W. (1994). Wissenserwerb mit logischen Bildern. In, B. Weidenmann (Hrsg.), *Wissenserwerb mit Bildern* (S. 95-147). Bern: Verlag Hans Huber.
- Schnotz, W. (2005). An Integrated Model of Text and Picture Comprehension. In R. E. Mayer (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (S. 49-69). Cambridge: Cambridge University Press.
- Weidenmann, B. (1994). *Lernen mit Bildmedien*. Weinheim: Beltz Psychologische Verlags Union.
- Weidenmann, B. (2001). Lernen mit Medien. In A. Krapp & B. Weidenmann (Hrsg.), *Pädagogische Psychologie* (S. 415-465). Weinheim: Beltz Psychologische Verlags Union.

Danksagung

Ich möchte mich bedanken bei...

... Prof. Dr. Damian Läge, für die ausgezeichnete Betreuung beim Erstellen dieser Arbeit.

... den Mitarbeitern des Projekt-Teams PTO (Damian Läge, Roland Streule, René Oberholzer und Samy Egli) für die spannende Zusammenarbeit.

... Angela Steiner, für die wertvolle Unterstützung und den moralischen Beistand.

... meinen Eltern, Armella und Felix Häne-Züger, welche mir diese Ausbildung ermöglicht und mich dabei immer unterstützt haben.

Curriculum Vitae

Personalien

Vorname, Name: Matthias Häne
Geschäftsadresse: Binzmühlestrasse 14 / Box 28, 8050 Zürich
Telefon G.: +41 (0)44 635 74 44
E-Mail: m.haene@psychologie.uzh.ch
URL: <http://www.pto.uzh.ch>
Geboren: 16.03.1977
Heimatort: Kirchberg (SG)
Zivilstand: ledig

Ausbildung

2005-2008 Doktorand an der Universität Zürich, Psychologisches Institut, Angewandte Kognitionspsychologie im SVC-Projekt „Psychopathology Taught Online (PTO)“.
Dissertationsschrift: Evaluation von Adaptivität im E-Learning: Qualitätssicherung und Qualitätskontrolle einer multimedialen Lernumgebung in der Psychopathologieausbildung.

1998-2004 Studium der Psychologie, Betriebswirtschaftslehre und Politikwissenschaft an der Universität Zürich.
Lizentiat: November 2004.
Lizentiatsarbeit: Untersuchung der Universalität und der Stabilität des semantischen Raums.

1992-1997 Kantonsschule Heerbrugg, E-Matura

1990-1992 Sekundarschule Rorschach

1984-1990 Primarschule Rorschach

Beruflicher Werdegang

| | |
|-----------------|---|
| 2006 – 2007 | Projektmitarbeiter (Evaluation) im Projekt „Psychopathology Taught Online“ (PTO) des Swiss Virtual Campus (SVC) Projekt-Nr. 3-008, an der Universität Zürich, Psychologisches Institut, Angewandte Kognitionspsychologie (http://www.pto.uzh.ch). |
| 2005-dato | Berufliche Tätigkeit als Consultant beim Usability-Dienstleistungsunternehmen soultank AG in Zug. |
| 10.2003-12.2003 | Praktikum beim Usability-Dienstleistungsunternehmen soultank AG in Zug. |
| 2002-2004 | 3 Semesterassistenzen am Psychologischen Institut der Universität Zürich, Abteilung Allgemeine Psychologie (unter anderem im Projekt zur Ästhetik in Zusammenarbeit mit der Universität Leipzig). |
| 02.2002-03.2002 | Praktikum bei Marketing, Research and Consulting (MRC) in Zug. |